

PATENT  
[MOBT:217]  
[38-21(15646)B US]

APPLICATION FOR UNITED STATES LETTERS PATENT  
for  
PREPARATION OF DEALLERGENIZED PROTEINS AND PERMUTEINS  
by  
*Murtaza F. Alibhai*  
*James D. Astwood*  
*Charles F. McWherter and*  
*Hugh A. Sampson*

EXPRESS MAIL MAILING LABEL	
NUMBER	<u>E15212 77946 U S</u>
DATE OF DEPOSIT	<u>1-5-01</u>
I hereby certify that this paper or fee is being deposited with the United States Postal Service "EXPRESS MAIL POST OFFICE TO ADDRESSEE" service under 37 C.F.R. 1.10 on the date indicated above and is addressed to: Assistant Commissioner for Patents, Washington D.C. 20231.	
<u>Elizabeth Gray</u> Signature	

1 FIELD OF THE INVENTION

2 The invention relates generally to non-naturally occurring novel proteins which  
3 have insecticidal properties, and more specifically to the design, preparation, and use of  
4 proteins that have been deallergenized while maintaining their insecticidal properties.  
5 Deallergenized patatin proteins include variants that have had allergenic sequences  
6 modified, and permuteins that have had their amino acid sequences rearranged at one or  
7 more breakpoints.

8 BACKGROUND OF THE INVENTION

9 Insecticidal proteins

10 The use of natural products, including proteins, is a well known method of  
11 controlling many insect, fungal, viral, bacterial, and nematode pathogens. For example,  
12 endotoxins of *Bacillus thuringiensis* (*B.t.*) are used to control both lepidopteran and  
13 coleopteran insect pests. Genes producing these endotoxins have been introduced into  
14 and expressed by various plants, including cotton, tobacco, and tomato. There are,  
15 however, several economically important insect pests such as boll weevil (BWV),  
16 *Anthonomus grandis*, and corn rootworm (CRW), *Diabrotica* spp. that are not as  
17 susceptible to *B.t.* endotoxins as are insects such as lepidopterans. In addition, having  
18 other, different gene products for control of insects which are susceptible to *B.t.*  
19 endotoxins is important, if not vital, for resistance management.

20 It has been recently discovered that the major storage protein of potato tubers,  
21 patatins (Gaillaire, T., *Biochem. J.* 121: 379-390, 1971; Racusen, D., *Can. J. Bot.*, 62:  
22 1640-1644, 1984; Andrews, D.L., *et al.*, *Biochem. J.*, 252: 199-206, 1988), will control  
23 various insects, including western rootworm (WCRW, *Diabrotica virgifera*), southern  
24 corn rootworm (SCRW, *Diabrotica undecimpunctata*), and boll weevil (BWV,  
25 *Anthonomus grandis*) (U.S. Patent No. 5,743,477). Patatins are lethal to some larvae and  
26 will stunt the growth of survivors so that maturation is prevented or severely delayed,  
27 resulting in no reproduction. These proteins, have nonspecific lipid acyl hydrolase  
28 activity and studies have shown that the enzyme activity is essential for its insecticidal

1 activity (Strickland, J.A., *et al.*, *Plant Physiol.*, 109: 667-674, 1995; U.S. Patent No.  
2 5,743,477). Patatins can be applied directly to the plants or introduced in other ways well  
3 known in the art, such as through the application of plant-colonizing microorganisms,  
4 which have been transformed to produce the enzymes, or by the plants themselves after  
5 similar transformation.

6 In potato, the patatins are found predominantly in tubers, but also at much lower  
7 levels in other plant organs (Hofgen, R. and Willmitzer, L., *Plant Science*, 66: 221-230,  
8 1990). Genes that encode patatins have been previously isolated by Mignery, G.A., *et al.*  
9 (*Nucleic Acids Research*, 12: 7987-8000, 1984; Mignery, G.A., *et al.*, *Gene*, 62: 27-44,  
10 1988; Stiekema, *et al.*, *Plant Mol. Biol.*, 11: 255-269, 1988) and others. Patatins are  
11 found in other plants, particularly solanaceous species (Ganal, *et al.*, *Mol. Gen. Genetics*,  
12 225: 501-509, 1991; Vancanneyt, *et al.*, *Plant Cell*, 1: 533-540, 1989) and recently Zea  
13 mays (WO 96/37615). Rosahl, *et al.* (*EMBO J.*, 6: 1155-1159, 1987) transferred it to  
14 tobacco plants, and observed expression of patatin, demonstrating that the patatin genes  
15 can be heterologously expressed by plants.

16 Patatin is an attractive for use in planta as an insect control agent, but  
17 unfortunately a small segment of the population displays allergic reactions to patatin  
18 proteins, and in particular to potato patatin, as described below.

19 **Food allergens**

20 There are a variety of proteins that cause allergic reactions. Proteins that have  
21 been identified as causing an allergic reaction in hypersensitive patients occur in many  
22 plant and animal derived foods, pollens, fungal spores, insect venoms, insect feces, and  
23 animal dander and urine (King, H.C., *Ear Nose Throat J.*, 73(4): 237-241, 1994 ;  
24 Astwood, J.D., *et al.*, *Clin. Exp. Allergy*, 25: 66-72, 1995; Astwood, J.D. and Fuchs R.L.,  
25 *Monographs in allergy Vol. 32: Highlights in food allergy*, pp. 105-120, 1996; Metcalfe,  
26 D.D., *et al.*, *Critical Reviews in Food Science and Nutrition*, 36S: 165-186, 1996 ). The  
27 offending proteins of many major sources of allergens have been characterized by clinical  
28 and molecular methods. The functions of allergenic proteins *in vivo* are diverse, ranging  
29 from enzymes to regulators of the cell cytoskeleton.

1 To understand the molecular basis of allergic disease, the important IgE binding  
2 epitopes of many allergen proteins have been mapped (Elsayed, S. and Apold, J., *Allergy*  
3 38(7): 449-459, 1983; Elsayed, S. et al., *Scand J. Clin. Lab. Invest. Suppl.* 204: 17-31  
4 1991; Zhang, L., et al., *Mol. Immunol.* 29(11): 1383-1389, 1992). The optimal peptide  
5 length for IgE binding has been reported to be between 8 and 12 amino acids.  
6 Conservation of epitope sequences is observed in homologous allergens of disparate  
7 species (Astwood, J.D., et al., *Clin. Exp. Allergy*, 25: 66-72, 1995). Indeed, conservative  
8 substitutions introduced by site-directed mutagenesis reduce IgE binding of known  
9 epitopes when presented as peptides.

10 Food allergy occurs in 2-6 % of the population. Eight foods or food groups (milk,  
11 eggs, fish, crustacea, wheat, peanuts, soybeans, and tree nuts) account for 90% of  
12 allergies to foods. Nevertheless, over 160 different foods have been reported to cause  
13 adverse reactions, including potato (Hefle, S., et al., *Crit. Rev. in Food Sci. Nutr.*, 36S:  
14 69-90, 1996).

15 Mode of action of allergens

16 Regardless of the identity of the allergen, it is theorized that the underlying  
17 mechanism of allergen response is the same. Immediate hypersensitivity (or anaphylactic  
18 response) is a form of allergic reaction which develops very quickly, i.e., within seconds  
19 or minutes of exposure of the patient to the causative allergen, and is mediated by B  
20 lymphocyte IgE antibody produciton. Allergic patients exhibit elevated levels of IgE,  
21 mediating hypersensitivity by priming mast cells which are abundant in the skin,  
22 lymphoid organs, in the membranes of the eye, nose and mouth, and in the respiratory  
23 tree and intestines. The IgE in allergy-suffering patients becomes bound to the IgE  
24 receptors of mast cells. When this bound IgE is subsequently contacted by the  
25 appropriate allergen, the mast cell is caused to degranulate and release various substances  
26 such as histamine into the surrounding tissue (Church et al. In: Kay, A.B. ed., *Allergy*  
27 and *Allergic Diseases*, Oxford, Blackwell Science, pp. 149-197, 1997).

28 It is the release of these substances which is responsible for the clinical symptoms  
29 typical of immediate hypersensitivity, namely contraction of smooth muscle in the  
30 airways or in the intestine, the dilation of small blood vessels, and the increase in their

1 permeability to water and plasma proteins, the secretion of thick sticky mucus, and (in the  
2 skin) the stimulation of nerve endings that result in itching or pain. Immediate  
3 hypersensitivity is, at best, a nuisance to the suffer; at worst it can present very serious  
4 problems and can in rare cases even result in death.

5 Allergic reactions to potato

6 Food allergy to potato is considered rare in the general population (Castells, M.C.,  
7 *et al.*, *Allergy Clin. Immunol.*, 8: 1110-1114, 1986; Hannuksela, M., *et al.*, *Contact  
8 Dermatitis*, 3: 79-84, 1977; Golbert, T.M., *et al.*, *Journal of Allergy*, 44: 96-107, 1969).  
9 Approximately 200 individuals have participated in published clinical accounts of potato  
10 allergy (Hefle, S. *et al.*, *Critical Reviews in Food Science and Nutrition*, 36S: 69-90,  
11 1996). A number of IgE binding proteins have been identified in potato tuber extracts  
12 (see Table 1), however the amino acid sequence and function of these proteins has not  
13 been determined (Wahl, R., *et al.*, *Intl. Arch. Allergy Appl. Immunol.*, 92: 168-174, 1990).

14 Table 1: Studies of potato tuber IgE-binding proteins (allergens)

Study	Protein Characteristics
(Castells, M.C. <i>et al.</i> <i>J. Allergy Clin. Immunol.</i> 78, 1110-1114, 1986)	Unknown 14 to 40 kDa
(Wahl, R. <i>et al.</i> <i>Int. Arch. Allergy Appl. Immunol.</i> 92: 168-174, 1990)	Unknown 42/43 kDa
	Unknown 65 kDa
	Unknown 26 kDa
	Unknown 20 kDa
	Unknown 14 kDa
	Unknown < 14 kDa (~ 5 kDa)
(Ebner, C. <i>et al.</i> in: Wuthrich, B. & Ortolani, C. (eds.), <i>Highlights in food allergy. Monographs in Allergy, Volume 32</i> Basil,Karger, pp. 73-77, 1996)	Unknown 42/43 kDa
	Unknown 23 kDa
	Unknown ~ 16 kDa
	Unknown < 14 kDa (~ 5 kDa)

15 Improved safety from the use of hypoallergenic proteins

16 Patatin has been identified as an allergenic protein (Seppala, U. *et al.*, *J. Allergy  
17 Clin. Immunol.* 103:165-171, 1999). Accordingly, potato allergic subjects may not be

1 able to safely consume products containing unmodified patatin protein, such as crops to  
2 which foliar applications of patatins have been applied, or crops which have been  
3 engineered to express patatin. In addition, proliferation of food allergens in the food  
4 supply is considered hazardous (Metcalfe, D.D., *et al.*, *Critical Reviews and Food*  
5 *Science and Nutrition*, 36S: 165-186, 1996). There are additional concerns regarding the  
6 use of potentially allergenic food proteins where workers might be exposed to airborne  
7 particulates, initiating a new allergic response (Moneret-Vautrin, D.A., *et al.*, *Rev. Med.*  
8 *Interne.*, 17(7): 551-557, 1996).

9 Permuteins

10 Novel proteins generated by the method of sequence transposition resembles that  
11 of naturally occurring pairs of proteins that are related by linear reorganization of their  
12 amino acid sequences (Cunningham, *et al.* *Proc. Natl. Sci., U.S.A.*, 76: 3218-3222, 1979;  
13 Teather, *et al.*, *J. Bacteriol.*, 172: 3837-3841, 1990; Schimming, *et al.*, *Eur. J. Biochem.*,  
14 204: 13-19, 1992; Yamiuchi, *et al.*, *FEBS Lett.*, 260: 127-130, 1991; MacGregor, *et al.*,  
15 *FEBS. Lett.*, 378: 263-266, 1996). The first *in vitro* application of sequence  
16 rearrangement to proteins was described by Goldenberg and Creighton (Goldenberg and  
17 Creighton, *J. Mol. Biol.*, 165: 407-413, 1983). A new N-terminus is selected at an  
18 internal site (breakpoint) of the original sequence, the new sequence having the same  
19 order of amino acids as the original from the breakpoint until it reaches an amino acid  
20 that is at or near the original C-terminus. At this point the new sequence is joined, either  
21 directly or through an additional portion or sequence (linker), to an amino acid that is at  
22 or near the original N-terminus, and the new sequence continues with the same sequence  
23 as the original until it reaches a point that is at or near the amino acid that was N-terminal  
24 to the breakpoint site of the original sequence, this residue forming the new C-terminus  
25 of the chain. This approach has been applied to proteins which range in size from 58 to  
26 462 amino acids and represent a broad range of structural classes (Goldenberg and  
27 Creighton, *J. Mol. Biol.*, 165: 407-413, 1983; Li and Coffino, *Mol. Cell. Biol.*, 13: 2377-  
28 2383, 1993; Zhang, *et al.*, *Nature Struct. Biol.*, 1: 434-438, 1995; Buchwalder, *et al.*,  
29 *Biochemistry*, 31: 1621-1630, 1994; Protasova, *et al.*, *Prot. Eng.*, 7: 1373-1377, 1995;  
30 Mullins, *et al.*, *J. Am. Chem. Soc.*, 116: 5529-5533, 1994; Garrett, *et al.*, *Protein Science*,

1 5: 204-211, 1996; Hahn, *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 91: 10417-10421, 1994;  
2 Yang and Schachman, *Proc. Natl. Acad. Sci. U.S.A.*, 90: 11980-11984, 1993; Luger, *et*  
3 *al.*, *Science*, 243: 206-210, 1989; Luger, *et al.*, *Prot. Eng.*, 3: 249-258, 1990; Lin, *et al.*,  
4 *Protein Science*, 4: 159-166, 1995; Vignais, *et al.*, *Protein Science*, 4: 994-1000, 1995;  
5 Ritco-Vonsovici, *et al.*, *Biochemistry*, 34: 16543-16551, 1995; Horlick, *et al.*, *Protein*  
6 *Eng.*, 5: 427-431, 1992; Kreitman, *et al.*, *Cytokine*, 7: 311-318, 1995; Viguera, *et al.*,  
7 *Mol. Biol.*, 247: 670-681, 1995; Koebnik and Kramer, *J. Mol. Biol.*, 250: 617-626, 1995;  
8 Kreitman, *et al.*, *Proc. Natl. Acad. Sci.*, 91: 6889-6893, 1994).

9 There exists a need for the development of plant expressible insecticidal proteins  
10 which possess minimal or no allergenic properties.

## 11 SUMMARY OF THE INVENTION

12 Novel protein sequences, and nucleic acid sequences encoding them are disclosed.  
13 The proteins maintain desirable enzymatic and insecticidal properties while displaying  
14 reduced or eliminated allergenicity.

15 Allergenic epitopes are identified by scanning overlapping peptide sequences with  
16 an immunoreactivity assay. Alanine scanning and 'rational substitution' is performed on  
17 identified peptide sequences to determine specific amino acids which contribute to  
18 antibody binding, and presumably, to the allergenic properties of the whole protein.  
19 Individual mutations are introduced into the whole protein sequence by methods such as  
20 site directed mutagenesis of the encoding nucleic acid sequence to delete or modify the  
21 allergenic sequences.

22 Glycosylation target residues are identified within amino acid sequences of  
23 proteins which have demonstrated allergy eliciting properties. Glycosylation target  
24 amino acid residues are rationally substituted with other amino acid residues to eliminate  
25 glycosylation and to provide a variant deglycosylated protein. The variant protein may  
26 then exhibit reduced allergen eliciting properties and may also exhibit reduced binding to  
27 IgE within serum of patients observed to be allergic to said glycosylated protein.

28 Permuteins of the deallergenized protein sequences can be constructed to further  
29 reduce or eliminate allergic reactions. The encoding nucleic acid sequence is modified to

1 produce a non-naturally occurring protein having a linear amino acid sequence different  
2 from the naturally occurring protein sequence, while maintaining enzymatic and  
3 insecticidal properties. The permutein is preferably produced in plant cells, and more  
4 preferably produced at a concentration which is toxic to insects ingesting the plant cells.

5 Methods for reducing, eliminating, or decreasing allergen eliciting properties of a  
6 protein are specifically contemplated herein. Such methods comprise steps including  
7 identifying one or more patients exhibiting an allergic sensitivity to an allergen eliciting  
8 protein and obtaining a sample of serum from the patient; exposing the patient serum to a  
9 first set of synthetic overlapping peptides which represent the allergen eliciting protein in  
10 order to identify such peptides which exhibit epitopes which bind to IgE present within  
11 the allergic patients' serum and wherein the IgE present in the serum has a specific  
12 affinity for the said allergen eliciting protein; producing a second set of peptides which  
13 are variant peptides based on the first set of peptides which were identified to bind  
14 specifically to IgE present in patient serum, wherein the second set variant peptides  
15 exhibit alanine scanning or rational scanning amino acid substitutions which exhibit  
16 reduced, decreased, or eliminated IgE binding when compared to the first set non-variant  
17 peptides, and wherein such substitutions which reduce, eliminate or decrease IgE binding  
18 are identified as result effective substitutions; and modifying the amino acid sequence of  
19 the allergen eliciting protein to contain one or more of said result effective substitutions,  
20 wherein the modified protein is a variant of the allergen eliciting protein which lacks  
21 allergen eliciting protein or exhibits reduced allergen eliciting properties, and wherein the  
22 variant of the allergen eliciting protein comprising one or more result effective  
23 substitutions exhibits reduced, decreased, or totally eliminated binding of IgE present  
24 within said patients' serum.

25 The novel proteins can be used in controlling insects, as nutritional supplements,  
26 in immunotherapy protocols, and in other potential applications. Transgenic plant cells  
27 and plants containing the encoding nucleic acid sequence can be particularly beneficial in  
28 the control of insects, and as a nutritional/immunotherapy material.

## DESCRIPTION OF THE FIGURES

2 The following figures form part of the present specification and are included to  
3 further demonstrate certain aspects of the present invention. The invention can be better  
4 understood by reference to one or more of these drawings in combination with the  
5 detailed description of specific embodiments presented herein.

Figure 1 illustrates the alignment of potato patatin PatA (acyl lipid hydrolase) with patatin (acyl lipid hydrolase) homologs and related amino acid sequences, the homologs and related sequences being from both dicot and monocot plant species.

9 Figure 2 illustrates IgE binding to overlapping peptide sequences.

10 Figure 3 illustrates construction of nucleic acid sequences encoding patatin  
11 permutein proteins, and in this figure for illustrative purposes a breakpoint at position 247  
12 is shown.

## DESCRIPTION OF THE SEQUENCE LISTINGS

14 The following description of the sequence listing forms part of the present  
15 specification and is included to further demonstrate certain aspects of the present  
16 invention. The invention can be better understood by reference to one or more of these  
17 sequences in combination with the detailed description of specific embodiments  
18 presented herein.

SEQ ID NO:1	DNA sequence encoding a patatin (acyl lipid hydrolase) protein
SEQ ID NO:2	potato patatin protein sequence
SEQ ID NO:3	thermal amplification primer
SEQ ID NO:4	thermal amplification primer
SEQ ID NO:5	thermal amplification product
SEQ ID NO:6	Pre-cleavage patatin protein produced in <i>Pichia pastoris</i>
SEQ ID NO:7	Post-cleavage patatin protein produced in <i>Pichia pastoris</i>
SEQ ID NO:8	Y106F mutagenic primer
SEQ ID NO:9	Y129F mutagenic primer
SEQ ID NO:10	Y185F mutagenic primer

SEQ ID NO:11	Y193F mutagenic primer
SEQ ID NO:12	Y185F and Y193F mutagenic primer
SEQ ID NO:13	Y270F mutagenic primer
SEQ ID NO:14	Y316F mutagenic primer
SEQ ID NO:15	Y362F mutagenic primer
SEQ ID NO:16-104	Peptide scan sequences of a patatin protein
SEQ ID NO:105-241	Alanine and rational scan sequences of selected patatin peptides
SEQ ID NO:242	thermal amplification primer 27
SEQ ID NO:243	thermal amplification primer 48
SEQ ID NO:244	thermal amplification primer 47
SEQ ID NO:245	thermal amplification primer 36
SEQ ID NO:246	pMON37402 sequence encoding permutein protein
SEQ ID NO:247	Permutein protein encoded from pMON37402 sequence
SEQ ID NO:248	thermal amplification primer 58
SEQ ID NO:249	thermal amplification primer 59
SEQ ID NO:250	pMON37405 sequence encoding permutein protein
SEQ ID NO:251	Permutein protein encoded by pMON37405 sequence
SEQ ID NO:252	thermal amplification primer 60
SEQ ID NO:253	thermal amplification primer 61
SEQ ID NO:254	pMON37406 sequence encoding permutein protein
SEQ ID NO:255	Permutein protein encoded by pMON37406 sequence
SEQ ID NO:256	thermal amplification primer 62
SEQ ID NO:257	thermal amplification primer 63
SEQ ID NO:258	pMON37407 sequence encoding permutein protein
SEQ ID NO:259	Permutein protein encoded by pMON37407 sequence
SEQ ID NO:260	thermal amplification primer 60
SEQ ID NO:261	thermal amplification primer 65
SEQ ID NO:262	pMON37408 sequence encoding permutein protein
SEQ ID NO:263	Permutein protein encoded by pMON37408 sequence
SEQ ID NO:264	pMON40701 sequence encoding permutein protein

SEQ ID NO:265	Permutein protein encoded by pMON40701 sequence
SEQ ID NO:266	thermal amplification primer Syn1
SEQ ID NO:267	thermal amplification primer Syn2
SEQ ID NO:268	thermal amplification primer Syn3
SEQ ID NO:269	thermal amplification primer Syn4
SEQ ID NO:270	pMON40703 sequence encoding permutein protein
SEQ ID NO:271	Permutein protein encoded by pMON40703 sequence
SEQ ID NO:272	thermal amplification primer Syn10
SEQ ID NO:273	thermal amplification primer Syn11
SEQ ID NO:274	pMON40705 sequence encoding permutein protein
SEQ ID NO:275	Permutein protein encoded by pMON40705 sequence
SEQ ID NO:276-277	Permutein linker sequences
SEQ ID NO:278	Patatin isozyme PatA+ (including signal peptide)
SEQ ID NO:279	Patatin isozyme PatB+ (including signal peptide)
SEQ ID NO:280	Patatin isozyme PatFm (mature protein lacking signal peptide)
SEQ ID NO:281	Patatin isozyme PatIm (mature protein lacking signal peptide)
SEQ ID NO:282	Patatin isozyme PatL+ (including signal peptide)
SEQ ID NO:283	Rational substitution peptide
SEQ ID NO:284	Corn homolog peptide
SEQ ID NO:285	patatin homolog Pat17 DNA coding sequence and amino acid translation
SEQ ID NO:286	patatin homolog Pat17 amino acid sequence
SEQ ID NO:287	dicot patatin homolog amino acid sequence pentin1_phb
SEQ ID NO:288	dicot patatin homolog amino acid sequence 5c9_phb
SEQ ID NO:289	maize patatin homolog amino acid sequence corn1_pep
SEQ ID NO:290	maize patatin homolog amino acid sequence corn2_pep
SEQ ID NO:291	maize patatin homolog amino acid sequence corn3_pep
SEQ ID NO:292	maize patatin homolog amino acid sequence corn4_pep
SEQ ID NO:293	maize patatin homolog amino acid sequence corn5_pep

## DEFINITIONS

2 The following definitions are provided in order to aid those skilled in the art in  
3 understanding the detailed description of the present invention. Some words and phrases  
4 may also be defined in other sections of the specification. No limitation should be placed  
5 on the definitions presented for the terms below, where other meanings are evidenced  
6 elsewhere in the specification in addition to those specified below.

7 "Allergen" refers to a biological or chemical substance that induces an allergic  
8 reaction or response. An allergic response can be an immunoglobulin E-mediated  
9 response.

10 Amino acid codes: A (Ala) = alanine; C (Cys) = cysteine; D (Asp) = aspartic acid;  
11 E (Glu) = glutamic acid; F (Phe) = phenylalanine; G (Gly) = glycine; H (His) = histidine;  
12 I (Ile) = isoleucine; K (Lys) = lysine; L (Leu) = leucine; M (Met) = methionine; N (Asn) =  
13 asparagine; P (Pro) = proline; Q (Gln) = glutamine; R (Arg) = arginine; S (Ser) = serine;  
14 T (Thr) = threonine; V (Val) = valine; W (Trp) = tryptophan; Y (Tyr) = tyrosine.

15 “Amplification: refers to increasing the number of copies of a desired molecule.

16 "Coding sequence", "open reading frame", and "structural sequence" refer to the  
17 region of continuous sequential nucleic acid base pair triplets encoding a protein,  
18 polypeptide, or peptide sequence.

19 "Codon" refers to a sequence of three nucleotides that specify a particular amino  
20 acid.

21 "Complementarity" refers to the specific binding of adenine to thymine (or uracil  
22 in RNA) and cytosine to guanine on opposite strands of DNA or RNA.

“Deallergenize” (render hypoallergenic) refers to the method of engineering or modifying a protein or the encoding DNA such that the protein has a reduced or eliminated ability to induce an allergic response with respect to the ability of the unmodified protein. A deallergenized protein can be referred to as being hypoallergenic. The degree of deallergenization of a protein can be measured *in vitro* by the reduced binding of IgE antibodies.

29 "DNA segment heterologous to the promoter region" means that the coding DNA  
30 segment does not exist in nature in the same gene with the promoter to which it is now  
31 attached.

1       “DNA segment” refers to a DNA molecule that has been isolated free of total  
2       genomic DNA of a particular species.

3       “Electroporation” refers to a method of introducing foreign DNA into cells that  
4       uses a brief, high voltage DC (direct current) charge to permeabilize the host cells,  
5       causing them to take up extra-chromosomal DNA.

6       “Encoding DNA” refers to chromosomal DNA, plasmid DNA, cDNA, or  
7       synthetic DNA which encodes any of the enzymes discussed herein.

8       “Endogenous” refers to materials originating from within an organism or cell.

9       “Endonuclease” refers to an enzyme that hydrolyzes double stranded DNA at  
10      internal locations.

11      “Epitope” refers to a region on an allergen that interacts with the cells of the  
12      immune system. Epitopes are often further defined by the type of antibody or cell with  
13      which they interact, e.g. if the region reacts with B-cells or antibodies (IgE), it is called a  
14      B-cell epitope.

15      “Exogenous” refers to materials originating from outside of an organism or cell.  
16      This typically applies to nucleic acid molecules used in producing transformed or  
17      transgenic host cells and plants.

18      “Expressibly coupled” and “expressibly linked” refer to a promoter or promoter  
19      region and a coding or structural sequence in such an orientation and distance that  
20      transcription of the coding or structural sequence can be directed by the promoter or  
21      promoter region.

22      “Expression” refers to the transcription of a gene to produce the corresponding  
23      mRNA and translation of this mRNA to produce the corresponding gene product, i.e., a  
24      peptide, polypeptide, or protein.

25      “Heterologous DNA” refers to DNA from a source different than that of the  
26      recipient cell.

27      “Homologous DNA” refers to DNA from the same source as that of the recipient  
28      cell.

29      “Identity” refers to the degree of similarity between two nucleic acid or protein  
30      sequences. An alignment of the two sequences is performed by a suitable computer  
31      program. A widely used and accepted computer program for performing sequence

1 alignments is CLUSTALW v1.6 (Thompson, et al. *Nucl. Acids Res.*, 22: 4673-4680,  
2 1994). The number of matching bases or amino acids is divided by the total number of  
3 bases or amino acids, and multiplied by 100 to obtain a percent identity. For example, if  
4 two 580 base pair sequences had 145 matched bases, they would be 25 percent identical.  
5 If the two compared sequences are of different lengths, the number of matches is divided  
6 by the shorter of the two lengths. For example, if there were 100 matched amino acids  
7 between 200 and a 400 amino acid proteins, they are 50 percent identical with respect to  
8 the shorter sequence. If the shorter sequence is less than 150 bases or 50 amino acids in  
9 length, the number of matches are divided by 150 (for nucleic acid bases) or 50 (for  
10 amino acids), and multiplied by 100 to obtain a percent identity.

11 “IgE” (Immunoglobulin E) refers to a specific class of immunoglobulin secreted  
12 by B cells. IgE binds to specific receptors on Mast cells. Interaction of an allergen with  
13 mast cell-bound IgE may trigger allergic symptoms.

14 “Immunotherapy” refers to any type of treatment that targets the immune system.  
15 Allergy immunotherapy is a treatment in which a progressively increasing dose of an  
16 allergen is given in order to induce an immune response characterized by tolerance to the  
17 antigen/allergen, also known as desensitization.

18 “*In vitro*” refers to “in the laboratory” and/or “outside of a living organism”.

19 “*In vivo*” refers to “in a living organism”.

20 “Insecticidal polypeptide” refers to a polypeptide having insecticidal properties  
21 that adversely affects the growth and development of insect pests.

22 “Monocot” refers to plants having a single cotyledon (the first leaf of the embryo  
23 of seed plants); examples include cereals such as maize, rice, wheat, oats, and barley.

24 “Multiple cloning site” refers to an artificially constructed collection of restriction  
25 enzyme sites in a vector that facilitates insertion of foreign DNA into the vector.

26 “Mutation” refers to any change or alteration in a nucleic acid sequence. Several  
27 types exist, including point, frame shift, splicing, and insertion/deletions.

28 “Native” refers to “naturally occurring in the same organism”. For example, a  
29 native promoter is the promoter naturally found operatively linked to a given coding  
30 sequence in an organism. A native protein is one naturally found in nature and untouched  
31 or not otherwise manipulated by the hand of man.

1       “Nucleic acid segment” is a nucleic acid molecule that has been isolated free of  
2 total genomic DNA of a particular species, or that has been synthesized. Included with  
3 the term “nucleic acid segment” are DNA segments, recombinant vectors, plasmids,  
4 cosmids, phagemids, phage, viruses, etcetera.

5       “Nucleic acid” refers to deoxyribonucleic acid (DNA) and ribonucleic acid  
6 (RNA).

7       Nucleic acid codes: A = adenosine; C = cytosine; G = guanosine; T = thymidine;  
8 N = equimolar A, C, G, and T; I = deoxyinosine; K = equimolar G and T; R = equimolar  
9 A and G; S = equimolar C and G; W = equimolar A and T; Y = equimolar C and T.

10       “Open reading frame (ORF)” refers to a region of DNA or RNA encoding a  
11 peptide, polypeptide, or protein or capable of being translated to protein, or a region of  
12 DNA capable of being transcribed into an RNA product.

13       “Plasmid” refers to a circular, extrachromosomal, self-replicating piece of DNA.

14       “Point mutation” refers to an alteration of a single nucleotide in a nucleic acid  
15 sequence.

16       “Polymerase chain reaction (PCR)” refers to an enzymatic technique to create  
17 multiple copies of one sequence of nucleic acid. Copies of DNA sequence are prepared  
18 by shuttling a DNA polymerase between two oligonucleotides. The basis of this  
19 amplification method is multiple cycles of temperature changes to denature, then re-  
20 anneal amplifiers, followed by extension to synthesize new DNA strands in the region  
21 located between the flanking amplifiers. Also known as thermal amplification.

22       “Probe” refers to a polynucleotide sequence which is complementary to a target  
23 polynucleotide sequence in the analyte. An antibody can also be used as a probe to detect  
24 the presence of an antigen. In that sense, the antigen binding domain of the antibody has  
25 some detectable affinity for the antigen and binds thereto. The binding of the antibody to  
26 the antigen can be measured by means known in the art, such as by chemiluminescence,  
27 phosphorescence, fluorescence, colorimetric chemical deposition at the site of binding, or  
28 otherwise.

29       “Promoter” or “promoter region” refers to a DNA sequence, usually found  
30 upstream (5') to a coding sequence, that controls expression of the coding sequence by  
31 controlling production of messenger RNA (mRNA) by providing the recognition site for

1 RNA polymerase and/or other factors necessary for start of transcription at the correct  
2 site. As contemplated herein, a promoter or promoter region includes variations of  
3 promoters derived by means of ligation to various regulatory sequences, random or  
4 controlled mutagenesis, and addition or duplication of enhancer sequences. The  
5 promoter region disclosed herein, and biologically functional equivalents thereof, are  
6 responsible for driving the transcription of coding sequences under their control when  
7 introduced into a host as part of a suitable recombinant vector, as demonstrated by its  
8 ability to produce mRNA.

9 "Recombinant DNA construct" or "recombinant vector" refers to any agent such  
10 as a plasmid, cosmid, virus, autonomously replicating sequence, phage, or linear or  
11 circular single-stranded or double-stranded DNA or RNA nucleotide sequence, derived  
12 from any source, capable of genomic integration or autonomous replication, comprising a  
13 DNA molecule in which one or more DNA sequences have been linked in a functionally  
14 operative manner. Such recombinant DNA constructs or vectors are capable of  
15 introducing a 5' regulatory sequence or promoter region and a DNA sequence for a  
16 selected gene product into a cell in such a manner that the DNA sequence is transcribed  
17 into a functional mRNA which is translated and therefore expressed. Recombinant DNA  
18 constructs or recombinant vectors can be constructed to be capable of expressing  
19 antisense RNAs, in order to inhibit translation of a specific RNA of interest.

20 "Recombinant proteins", also referred to as "heterologous proteins", are proteins  
21 which are normally not produced by the host cell.

22 "Regeneration" refers to the process of growing a plant from a plant cell (e.g.,  
23 plant protoplast or explant).

24 "Regeneration" refers to the process of growing a plant from a plant cell (e.g.,  
25 plant protoplast or explant).

26 "Regulatory sequence" refers to a nucleotide sequence located upstream (5'),  
27 within, and/or downstream (3') to a DNA sequence encoding a selected gene product  
28 whose transcription and expression is controlled by the regulatory sequence in  
29 conjunction with the protein synthetic apparatus of the cell.

1        “Restriction enzyme” refers to an enzyme that recognizes a specific palindromic  
2 sequence of nucleotides in double stranded DNA and cleaves both strands; also called a  
3 restriction endonuclease. Cleavage typically occurs within the restriction site.

4        “Result-effective substitution” (RES) refers to an amino acid substitution within  
5 an IgE-binding region (epitope) of a target protein which reduces or eliminates the IgE  
6 binding by that epitope. Examples herein are directed to patatin protein and homologues,  
7 however, as will be readily recognized by those skilled in the art, the method is more  
8 broadly applicable to proteins other than patatins, and in particular is applicable to any  
9 protein exhibiting allergen eliciting properties.

10       “Selectable marker” refers to a nucleic acid sequence whose expression confers a  
11 phenotype facilitating identification of cells containing the nucleic acid sequence.  
12 Selectable markers include those which confer resistance to toxic chemicals (e.g.  
13 ampicillin resistance, kanamycin resistance), complement a nutritional deficiency (e.g.  
14 uracil, histidine, leucine), or impart a visually distinguishing characteristic (e.g. color  
15 changes or fluorescence).

16       “Transcription” refers to the process of producing an RNA copy from a DNA  
17 template.

18       “Transformation” refers to a process of introducing an exogenous nucleic acid  
19 sequence (e.g., a vector, recombinant nucleic acid molecule) into a cell or protoplast in  
20 which that exogenous nucleic acid is incorporated into a chromosome or is capable of  
21 autonomous replication.

22       “Transformed cell” is a cell whose DNA has been altered by the introduction of  
23 an exogenous nucleic acid molecule into that cell.

24       “Transgenic cell” refers to any cell derived from or regenerated from a  
25 transformed cell or derived from a transgenic cell. Exemplary transgenic cells include  
26 plant calli derived from a transformed plant cell and particular cells such as leaf, root,  
27 stem, *e.g.*, somatic cells, or reproductive (germ) cells obtained from a transgenic plant.

28       “Transgenic plant” refers to a plant or progeny thereof derived from a transformed  
29 plant cell or protoplast, wherein the plant DNA contains an introduced exogenous nucleic  
30 acid sequence not originally present in a native, non-transgenic plant of the same species.

1 Alternatively, the plant DNA can contain the introduced nucleic acid sequence in a higher  
2 copy number than in the native, non-transgenic plant of the same species.

3 "Translation" refers to the production of protein from messenger RNA.

4 "Vector" refers to a plasmid, cosmid, bacteriophage, or virus that carries foreign  
5 DNA into a host organism.

6 "Western blot" refers to protein or proteins that have been separated by  
7 electrophoresis, transferred and immobilized onto a solid support, then probed with an  
8 antibody.

9 **DETAILED DESCRIPTION OF THE INVENTION**

10 **Design of deallergenized patatin proteins**

11 Deallergenizing a protein can be accomplished by the identification of allergenic  
12 sites, followed by modification of the sites to reduce or eliminate the binding of  
13 antibodies to the sites. The IgE-binding regions of patatin were previously unreported.  
14 Mapping of the IgE epitopes was accomplished by synthesizing 10-mer peptides based on  
15 the patatin 17 protein sequence (SEQ ID NO: 2) which overlap by six amino acids. As  
16 potato proteins are denatured upon cooking potato products, it is expected that the 10-mer  
17 peptides sufficiently mimic the unfolded full length protein for antibody binding  
18 purposes. Peptides were identified based upon their ability to bind to IgE antibodies.  
19 Individual amino acids within the identified peptides were changed to reduce or eliminate  
20 binding to IgE present in sera from potato sensitive patients. These changes are termed  
21 result-effective amino acid substitutions (RES). The RES can be subsequently introduced  
22 into the full length protein by site directed mutagenesis of the encoding nucleic acid  
23 sequence or other means known in the art. Similar strategies have been employed  
24 elsewhere to determine the dominant IgE epitopes in a major peanut allergen (Stanley,  
25 J.S., *et al.*, *Arch. Biochem. Biophys.*, 342(2): 244-253, 1997).

26 Certain amino acid residues important for allergenicity of patatin are identified.  
27 Some of the designed patatin peptides wherein single amino acid residues were replaced  
28 with alanine or phenylalanine, showed significantly reduced or no binding to sera from  
29 potato sensitive patients.

1        A “deallergenized patatin” refers to a patatin protein differing in at least one of  
2        the amino acid residues as defined by the result effective substitutions resulting in the  
3        patatin protein having reduced reactivity towards sera from potato sensitive patients. The  
4        deallergenized patatin preferably maintains insecticidal properties, and preferably  
5        maintains its characteristic enzymatic profile.

6        Summary of method to deallergenize a patatin protein

7            ▪ Mapping of IgE epitopes by immunoassay of synthetic overlapping peptides  
8            using sera from potato sensitive patients;

9            ▪ Identification of result-effective substitutions by alanine scanning and/or  
10           rational scanning;

11           ▪ Modification of the amino acid sequence of patatin by site-directed  
12           mutagenesis of the encoding nucleic acid sequence;

13           ▪ Evaluation of enzyme activity (esterase) and/or insecticidal activity of the  
14           modified protein(s); and

15           ▪ Evaluation of the new protein(s) for allergenicity by IgE immunoassay.

16  
17        Nucleic acid sequences encoding patatin have been cloned by several  
18        investigators (e.g. Mignery, *et al.*, *Nucleic Acids Research*, 12: 7987-8000, 1984;  
19        Mignery, *et al.*, *Gene*, 62: 27-44, 1988; WO 94/21805; Canadian Patent Application No.  
20        2090552). These nucleic acid sequences can then be manipulated using site directed  
21        mutagenesis to encode a hypoallergenic patatin. These nucleic acid sequences can than  
22        be used to transform bacterial, yeast or plant cells, resulting in the production of  
23        hypoallergenic patatin protein.

24        Deallergenized patatin proteins

25        For simplicity, individual amino acids are referred to by their single letter codes.  
26        Correlation between the single letter codes, three letter codes, and full amino acid names  
27        is presented in the definitions section above.

28        One embodiment of the invention is an isolated deallergenized patatin protein.  
29        The protein is modified relative to the wild-type protein sequence such that they exhibit

1 reduced binding to anti-patatin antibodies such as those obtained from humans or animals  
2 allergic to potatoes. The reduced binding is measured relative to the binding of the  
3 unmodified patatin protein to the anti-patatin antibodies.

4 The deallergenized patatin protein can comprise SEQ ID NO:2 modified in one or  
5 more of the following regions, or SEQ ID NO:7 modified in one or more of the following  
6 regions. The single or multiple amino acid modifications reduce the binding of the  
7 modified protein relative to the binding of the corresponding unmodified protein. The  
8 regions for modification include amino acid positions 104-113 of SEQ ID NO:2 (85-94  
9 of SEQ ID NO:7), 128-137 of SEQ ID NO:2 (109-118 of SEQ ID NO:7), 184-197 of  
10 SEQ ID NO:2 (165-178 of SEQ ID NO:7), 264-277 of SEQ ID NO:2 (245-258 of SEQ  
11 ID NO:7), 316-325 of SEQ ID NO:2 (297-306 of SEQ ID NO:7), and 360-377 of SEQ ID  
12 NO:2 (341-358 of SEQ ID NO:7). The possible amino acid modifications include  
13 replacing an amino acid with A, E, F, P, or S. The modifications replace one or more  
14 amino acids in the identified regions, without increasing or decreasing the total number of  
15 amino acids in the protein.

16 Preferably, the deallergenized patatin protein comprises SEQ ID NO:2 modified  
17 by one or more changes, or SEQ ID NO:7 modified by one or more changes. SEQ ID  
18 NO:7 differs from wild type SEQ ID NO:2 in that the first 22 amino acids of SEQ ID  
19 NO:2 are replaced with EAE (Glu-Ala-Glu). For example, the changes to SEQ ID NO:2  
20 or SEQ ID NO:7 can be: the Y corresponding to position 106 of SEQ ID NO:2 or  
21 position 87 of SEQ ID NO:7 is replaced with F or A; the I corresponding to position 113  
22 of SEQ ID NO:2 or position 94 of SEQ ID NO:7 is replaced with A; the Y corresponding  
23 to position 129 of SEQ ID NO:2 or position 110 of SEQ ID NO:7 is replaced with F or  
24 A; the K corresponding to position 137 of SEQ ID NO:2 or position 118 of SEQ ID NO:7  
25 is replaced with A; the S corresponding to position 184 of SEQ ID NO:2 or position 165  
26 of SEQ ID NO:7 is replaced with A; the Y corresponding to position 185 of SEQ ID  
27 NO:2 or position 166 of SEQ ID NO:7 is replaced with F or A; the A corresponding to  
28 position 188 of SEQ ID NO:2 or position 169 of SEQ ID NO:7 is replaced with S; the T  
29 corresponding to position 192 of SEQ ID NO:2 or position 173 of SEQ ID NO:7 is  
30 replaced with A or P; the Y corresponding to position 193 of SEQ ID NO:2 or position  
31 174 of SEQ ID NO:7 is replaced with F or A; the K corresponding to position 268 of

1 SEQ ID NO:2 or position 249 of SEQ ID NO:7 is replaced with A or E; the T  
2 corresponding to position 269 of SEQ ID NO:2 or position 250 of SEQ ID NO:7 is  
3 replaced with A; the Y corresponding to position 270 of SEQ ID NO:2 or position 251 of  
4 SEQ ID NO:7 is replaced with F or A; the K corresponding to position 273 of SEQ ID  
5 NO:2 or position 254 of SEQ ID NO:7 is replaced with A; the K corresponding to  
6 position 313 of SEQ ID NO:2 or position 294 of SEQ ID NO:7 is replaced with E; the N  
7 corresponding to position 314 of SEQ ID NO:2 or position 295 of SEQ ID NO:7 is  
8 replaced with A; the N corresponding to position 315 of SEQ ID NO:2 or position 296 of  
9 SEQ ID NO:7 is replaced with A; the Y corresponding to position 316 of SEQ ID NO:2  
10 or position 297 of SEQ ID NO:7 is replaced with F or A; the Y corresponding to position  
11 362 of SEQ ID NO:2 or position 343 of SEQ ID NO:7 is replaced with F; the K  
12 corresponding to position 367 of SEQ ID NO:2 or position 348 of SEQ ID NO:7 is  
13 replaced with A; the R corresponding to position 368 of SEQ ID NO:2 or position 349 of  
14 SEQ ID NO:7 is replaced with A; the F corresponding to position 369 of SEQ ID NO:2  
15 or position 350 of SEQ ID NO:7 is replaced with A; the K corresponding to position 371  
16 of SEQ ID NO:2 or position 352 of SEQ ID NO:7 is replaced with A; the L  
17 corresponding to position 372 of SEQ ID NO:2 or position 353 of SEQ ID NO:7 is  
18 replaced with A; and the L corresponding to position 373 of SEQ ID NO:2 or position  
19 354 of SEQ ID NO:7 is replaced with A.

20 More preferably, SEQ ID NO:2 is modified by the following changes or SEQ ID  
21 NO:7 is modified by the following changes: the Y corresponding to position 106 of SEQ  
22 ID NO:2 or position 87 of SEQ ID NO:7 is replaced with F; the Y corresponding to  
23 position 129 of SEQ ID NO:2 or position 110 of SEQ ID NO:7 is replaced with F; the Y  
24 corresponding to position 185 of SEQ ID NO:2 or position 166 of SEQ ID NO:7 is  
25 replaced with F; the Y corresponding to position 193 of SEQ ID NO:2 or position 174 of  
26 SEQ ID NO:7 is replaced with F; the Y corresponding to position 270 of SEQ ID NO:2  
27 or position 251 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 316  
28 of SEQ ID NO:2 or position 297 of SEQ ID NO:7 is replaced with F; and the Y  
29 corresponding to position 362 of SEQ ID NO:2 or position 343 of SEQ ID NO:7 is  
30 replaced with F.

1           Most preferably, SEQ ID NO:2 is modified by the following changes or SEQ ID  
2   NO:7 is modified by the following changes: the Y corresponding to position 185 of SEQ  
3   ID NO:2 or position 166 of SEQ ID NO:7 is replaced with F; the Y corresponding to  
4   position 193 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is replaced with F; the Y  
5   corresponding to position 270 of SEQ ID NO:2 or position 251 of SEQ ID NO:7 is  
6   replaced with F; the Y corresponding to position 316 of SEQ ID NO:2 or position 297 of  
7   SEQ ID NO:7 is replaced with F; and the Y corresponding to position 362 of SEQ ID  
8   NO:2 or position 343 of SEQ ID NO:7 is replaced with F.

9           Nucleic acids

10          An additional embodiment of the invention is an isolated nucleic acid molecule  
11   segment comprising a structural nucleic acid sequence which encodes a deallergenized  
12   patatin protein.

13          The structural nucleic acid sequence can generally encode any deallergenized  
14   patatin protein. The structural nucleic acid sequence preferably encodes a deallergenized  
15   patatin protein comprising SEQ ID NO:2 modified in one or more of the following  
16   regions, or SEQ ID NO:7 modified in one or more of the following regions. The single  
17   or multiple amino acid modifications reduce the binding of the modified protein relative  
18   to the binding of the corresponding unmodified protein. The regions for modification  
19   include amino acid positions 104-113 of SEQ ID NO:2 (85-94 of SEQ ID NO:7), 128-  
20   137 of SEQ ID NO:2 (109-118 of SEQ ID NO:7), 184-197 of SEQ ID NO:2 (165-178 of  
21   SEQ ID NO:7), 264-277 of SEQ ID NO:2 (245-258 of SEQ ID NO:7), 316-325 of SEQ  
22   ID NO:2 (297-306 of SEQ ID NO:7), and 360-377 of SEQ ID NO:2 (341-358 of SEQ ID  
23   NO:7). The possible amino acid modifications include replacing an amino acid with A,  
24   E, F, P, or S. The modifications replace one or more amino acids in the identified  
25   regions, without increasing or decreasing the total number of amino acids in the protein.

26          Alternatively, the structural nucleic acid sequence encodes SEQ ID NO:2  
27   modified by one or more of the following changes or encoding SEQ ID NO:7 modified  
28   by one or more of the following changes: the Y corresponding to position 106 of SEQ ID  
29   NO:2 or position 87 of SEQ ID NO:7 is replaced with F or A; the I corresponding to  
30   position 113 of SEQ ID NO:2 or position 94 of SEQ ID NO:7 is replaced with A; the Y

1 corresponding to position 129 of SEQ ID NO:2 or position 110 of SEQ ID NO:7 is  
2 replaced with F or A; the K corresponding to position 137 of SEQ ID NO:2 or position  
3 118 of SEQ ID NO:7 is replaced with A; the S corresponding to position 184 of SEQ ID  
4 NO:2 or position 165 of SEQ ID NO:7 is replaced with A; the Y corresponding to  
5 position 185 of SEQ ID NO:2 or position 166 of SEQ ID NO:7 is replaced with F or A;  
6 the A corresponding to position 188 of SEQ ID NO:2 or position 169 of SEQ ID NO:7 is  
7 replaced with S; the T corresponding to position 192 of SEQ ID NO:2 or position 173 of  
8 SEQ ID NO:7 is replaced with A or P; the Y corresponding to position 193 of SEQ ID  
9 NO:2 or position 174 of SEQ ID NO:7 is replaced with F or A; the K corresponding to  
10 position 268 of SEQ ID NO:2 or position 249 of SEQ ID NO:7 is replaced with A or E;  
11 the T corresponding to position 269 of SEQ ID NO:2 or position 250 of SEQ ID NO:7 is  
12 replaced with A; the Y corresponding to position 270 of SEQ ID NO:2 or position 251 of  
13 SEQ ID NO:7 is replaced with F or A; the K corresponding to position 273 of SEQ ID  
14 NO:2 or position 254 of SEQ ID NO:7 is replaced with A; the K corresponding to  
15 position 313 of SEQ ID NO:2 or position 294 of SEQ ID NO:7 is replaced with E; the N  
16 corresponding to position 314 of SEQ ID NO:2 or position 295 of SEQ ID NO:7 is  
17 replaced with A; the N corresponding to position 315 of SEQ ID NO:2 or position 296 of  
18 SEQ ID NO:7 is replaced with A; the Y corresponding to position 316 of SEQ ID NO:2  
19 or position 297 of SEQ ID NO:7 is replaced with F or A; the Y corresponding to position  
20 362 of SEQ ID NO:2 or position 343 of SEQ ID NO:7 is replaced with F; the K  
21 corresponding to position 367 of SEQ ID NO:2 or position 348 of SEQ ID NO:7 is  
22 replaced with A; the R corresponding to position 368 of SEQ ID NO:2 or position 349 of  
23 SEQ ID NO:7 is replaced with A; the F corresponding to position 369 of SEQ ID NO:2  
24 or position 350 of SEQ ID NO:7 is replaced with A; the K corresponding to position 371  
25 of SEQ ID NO:2 or position 352 of SEQ ID NO:7 is replaced with A; the L  
26 corresponding to position 372 of SEQ ID NO:2 or position 353 of SEQ ID NO:7 is  
27 replaced with A; and the L corresponding to position 373 of SEQ ID NO:2 or position  
28 354 of SEQ ID NO:7 is replaced with A.

29 More preferably, the structural nucleic acid sequence encodes SEQ ID NO:2  
30 modified by the following changes or SEQ ID NO:7 modified by the following changes:  
31 the Y corresponding to position 106 of SEQ ID NO:2 or position 87 of SEQ ID NO:7 is

1 replaced with F; the Y corresponding to position 129 of SEQ ID NO:2 or position 110 of  
2 SEQ ID NO:7 is replaced with F; the Y corresponding to position 185 of SEQ ID NO:2  
3 or position 166 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 193  
4 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is replaced with F; the Y  
5 corresponding to position 270 of SEQ ID NO:2 or position 251 of SEQ ID NO:7 is  
6 replaced with F; the Y corresponding to position 316 of SEQ ID NO:2 or position 297 of  
7 SEQ ID NO:7 is replaced with F; and the Y corresponding to position 362 of SEQ ID  
8 NO:2 or position 343 of SEQ ID NO:7 is replaced with F.

9 Most preferably, the structural nucleic acid sequence encodes SEQ ID NO:2  
10 modified by the following changes or SEQ ID NO:7 modified by the following changes:  
11 the Y corresponding to position 185 of SEQ ID NO:2 or position 166 of SEQ ID NO:7 is  
12 replaced with F; the Y corresponding to position 193 of SEQ ID NO:2 or position 174 of  
13 SEQ ID NO:7 is replaced with F; the Y corresponding to position 270 of SEQ ID NO:2  
14 or position 251 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 316  
15 of SEQ ID NO:2 or position 297 of SEQ ID NO:7 is replaced with F; and the Y  
16 corresponding to position 362 of SEQ ID NO:2 or position 343 of SEQ ID NO:7 is  
17 replaced with F.

18 Recombinant vectors

19 An additional embodiment is directed towards recombinant vectors comprising a  
20 structural nucleic acid sequence which encodes a deallergenized patatin protein. The  
21 recombinant vector comprises operatively linked in the 5' to 3' orientation: a promoter  
22 that directs transcription of a structural nucleic acid sequence; a structural nucleic acid  
23 sequence, and a 3' transcription terminator.

24 The structural nucleic acid sequence can encode SEQ ID NO:2 modified in one or  
25 more of the following regions, or SEQ ID NO:7 modified in one or more of the following  
26 regions. The single or multiple amino acid modifications reduce the binding of the  
27 modified protein relative to the binding of the corresponding unmodified protein. The  
28 regions for modification include amino acid positions 104-113 of SEQ ID NO:2 (85-94  
29 of SEQ ID NO:7), 128-137 of SEQ ID NO:2 (109-118 of SEQ ID NO:7), 184-197 of  
30 SEQ ID NO:2 (165-178 of SEQ ID NO:7), 264-277 of SEQ ID NO:2 (245-258 of SEQ

1 ID NO:7), 316-325 of SEQ ID NO:2 (297-306 of SEQ ID NO:7), and 360-377 of SEQ ID  
2 NO:2 (341-358 of SEQ ID NO:7). The possible amino acid modifications include  
3 replacing an amino acid with A, E, F, P, or S. The modifications replace one or more  
4 amino acids in the identified regions, without increasing or decreasing the total number of  
5 amino acids in the protein.

6 Alternatively, the recombinant vector comprises operatively linked in the 5' to 3'  
7 orientation: a promoter that directs transcription of a structural nucleic acid sequence; a  
8 structural nucleic acid sequence encoding SEQ ID NO:2 modified by one or more of the  
9 following changes or encoding SEQ ID NO:7 modified by one or more of the following  
10 changes: the Y corresponding to position 106 of SEQ ID NO:2 or position 87 of SEQ ID  
11 NO:7 is replaced with F or A; the I corresponding to position 113 of SEQ ID NO:2 or  
12 position 94 of SEQ ID NO:7 is replaced with A; the Y corresponding to position 129 of  
13 SEQ ID NO:2 or position 110 of SEQ ID NO:7 is replaced with F or A; the K  
14 corresponding to position 137 of SEQ ID NO:2 or position 118 of SEQ ID NO:7 is  
15 replaced with A; the S corresponding to position 184 of SEQ ID NO:2 or position 165 of  
16 SEQ ID NO:7 is replaced with A; the Y corresponding to position 185 of SEQ ID NO:2  
17 or position 166 of SEQ ID NO:7 is replaced with F or A; the A corresponding to position  
18 188 of SEQ ID NO:2 or position 169 of SEQ ID NO:7 is replaced with S; the T  
19 corresponding to position 192 of SEQ ID NO:2 or position 173 of SEQ ID NO:7 is  
20 replaced with A or P; the Y corresponding to position 193 of SEQ ID NO:2 or position  
21 174 of SEQ ID NO:7 is replaced with F or A; the K corresponding to position 268 of  
22 SEQ ID NO:2 or position 249 of SEQ ID NO:7 is replaced with A or E; the T  
23 corresponding to position 269 of SEQ ID NO:2 or position 250 of SEQ ID NO:7 is  
24 replaced with A; the Y corresponding to position 270 of SEQ ID NO:2 or position 251 of  
25 SEQ ID NO:7 is replaced with F or A; the K corresponding to position 273 of SEQ ID  
26 NO:2 or position 254 of SEQ ID NO:7 is replaced with A; the K corresponding to  
27 position 313 of SEQ ID NO:2 or position 294 of SEQ ID NO:7 is replaced with E; the N  
28 corresponding to position 314 of SEQ ID NO:2 or position 295 of SEQ ID NO:7 is  
29 replaced with A; the N corresponding to position 315 of SEQ ID NO:2 or position 296 of  
30 SEQ ID NO:7 is replaced with A; the Y corresponding to position 316 of SEQ ID NO:2  
31 or position 297 of SEQ ID NO:7 is replaced with F or A; the Y corresponding to position

1 362 of SEQ ID NO:2 or position 343 of SEQ ID NO:7 is replaced with F; the K  
2 corresponding to position 367 of SEQ ID NO:2 or position 348 of SEQ ID NO:7 is  
3 replaced with A; the R corresponding to position 368 of SEQ ID NO:2 or position 349 of  
4 SEQ ID NO:7 is replaced with A; the F corresponding to position 369 of SEQ ID NO:2  
5 or position 350 of SEQ ID NO:7 is replaced with A; the K corresponding to position 371  
6 of SEQ ID NO:2 or position 352 of SEQ ID NO:7 is replaced with A; the L  
7 corresponding to position 372 of SEQ ID NO:2 or position 353 of SEQ ID NO:7 is  
8 replaced with A; and the L corresponding to position 373 of SEQ ID NO:2 or position  
9 354 of SEQ ID NO:7 is replaced with A; and a 3' transcription terminator.

10 More preferably, the vector comprises a structural nucleic acid sequence encoding  
11 SEQ ID NO:2 modified by the following changes or SEQ ID NO:7 modified by the  
12 following changes: the Y corresponding to position 106 of SEQ ID NO:2 or position 87  
13 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 129 of SEQ ID  
14 NO:2 or position 110 of SEQ ID NO:7 is replaced with F; the Y corresponding to  
15 position 185 of SEQ ID NO:2 or position 166 of SEQ ID NO:7 is replaced with F; the Y  
16 corresponding to position 193 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is  
17 replaced with F; the Y corresponding to position 270 of SEQ ID NO:2 or position 251 of  
18 SEQ ID NO:7 is replaced with F; the Y corresponding to position 316 of SEQ ID NO:2  
19 or position 297 of SEQ ID NO:7 is replaced with F; and the Y corresponding to position  
20 362 of SEQ ID NO:2 or position 343 of SEQ ID NO:7 is replaced with F.

21 Most preferably, the vector comprises a structural nucleic acid sequence encoding  
22 SEQ ID NO:2 modified by the following changes or SEQ ID NO:7 modified by the  
23 following changes: the Y corresponding to position 185 of SEQ ID NO:2 or position 166  
24 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 193 of SEQ ID  
25 NO:2 or position 174 of SEQ ID NO:7 is replaced with F; the Y corresponding to  
26 position 270 of SEQ ID NO:2 or position 251 of SEQ ID NO:7 is replaced with F; the Y  
27 corresponding to position 316 of SEQ ID NO:2 or position 297 of SEQ ID NO:7 is  
28 replaced with F; and the Y corresponding to position 362 of SEQ ID NO:2 or position  
29 343 of SEQ ID NO:7 is replaced with F.

1                   Recombinant host cells

2                   A further embodiment of the invention is directed towards recombinant host cells  
3                   comprising a structural nucleic acid sequence encoding a deallergenized patatin protein.  
4                   The recombinant host cell preferably produces a deallergenized patatin protein. More  
5                   preferably, the recombinant host cell produces a deallergenized patatin protein in a  
6                   concentration sufficient to inhibit growth or to kill an insect which ingests the  
7                   recombinant host cell. The recombinant host cell can generally comprise any structural  
8                   nucleic acid sequence encoding a deallergenized patatin protein.

9                   The recombinant host cell can comprise a structural nucleic acid sequence  
10                  encoding SEQ ID NO:2 modified in one or more of the following regions, or SEQ ID  
11                  NO:7 modified in one or more of the following regions. The single or multiple amino  
12                  acid modifications reduce the binding of the modified protein relative to the binding of  
13                  the corresponding unmodified protein. The regions for modification include amino acid  
14                  positions 104-113 of SEQ ID NO:2 (85-94 of SEQ ID NO:7), 128-137 of SEQ ID NO:2  
15                  (109-118 of SEQ ID NO:7), 184-197 of SEQ ID NO:2 (165-178 of SEQ ID NO:7), 264-  
16                  277 of SEQ ID NO:2 (245-258 of SEQ ID NO:7), 316-325 of SEQ ID NO:2 (297-306 of  
17                  SEQ ID NO:7), and 360-377 of SEQ ID NO:2 (341-358 of SEQ ID NO:7). The possible  
18                  amino acid modifications include replacing an amino acid with A, E, F, P, or S. The  
19                  modifications replace one or more amino acids in the identified regions, without  
20                  increasing or decreasing the total number of amino acids in the protein.

21                  Alternatively, the recombinant host cell comprises a structural nucleic acid  
22                  sequence encoding SEQ ID NO:2 modified by one or more of the following changes or  
23                  encoding SEQ ID NO:7 modified by one or more of the following changes: the Y  
24                  corresponding to position 106 of SEQ ID NO:2 or position 87 of SEQ ID NO:7 is  
25                  replaced with F or A; the I corresponding to position 113 of SEQ ID NO:2 or position 94  
26                  of SEQ ID NO:7 is replaced with A; the Y corresponding to position 129 of SEQ ID  
27                  NO:2 or position 110 of SEQ ID NO:7 is replaced with F or A; the K corresponding to  
28                  position 137 of SEQ ID NO:2 or position 118 of SEQ ID NO:7 is replaced with A; the S  
29                  corresponding to position 184 of SEQ ID NO:2 or position 165 of SEQ ID NO:7 is  
30                  replaced with A; the Y corresponding to position 185 of SEQ ID NO:2 or position 166 of  
31                  SEQ ID NO:7 is replaced with F or A; the A corresponding to position 188 of SEQ ID

1 NO:2 or position 169 of SEQ ID NO:7 is replaced with S; the T corresponding to position  
2 192 of SEQ ID NO:2 or position 173 of SEQ ID NO:7 is replaced with A or P; the Y  
3 corresponding to position 193 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is  
4 replaced with F or A; the K corresponding to position 268 of SEQ ID NO:2 or position  
5 249 of SEQ ID NO:7 is replaced with A or E; the T corresponding to position 269 of  
6 SEQ ID NO:2 or position 250 of SEQ ID NO:7 is replaced with A; the Y corresponding  
7 to position 270 of SEQ ID NO:2 or position 251 of SEQ ID NO:7 is replaced with F or  
8 A; the K corresponding to position 273 of SEQ ID NO:2 or position 254 of SEQ ID NO:7  
9 is replaced with A; the K corresponding to position 313 of SEQ ID NO:2 or position 294  
10 of SEQ ID NO:7 is replaced with E; the N corresponding to position 314 of SEQ ID  
11 NO:2 or position 295 of SEQ ID NO:7 is replaced with A; the N corresponding to  
12 position 315 of SEQ ID NO:2 or position 296 of SEQ ID NO:7 is replaced with A; the Y  
13 corresponding to position 316 of SEQ ID NO:2 or position 297 of SEQ ID NO:7 is  
14 replaced with F or A; the Y corresponding to position 362 of SEQ ID NO:2 or position  
15 343 of SEQ ID NO:7 is replaced with F; the K corresponding to position 367 of SEQ ID  
16 NO:2 or position 348 of SEQ ID NO:7 is replaced with A; the R corresponding to  
17 position 368 of SEQ ID NO:2 or position 349 of SEQ ID NO:7 is replaced with A; the F  
18 corresponding to position 369 of SEQ ID NO:2 or position 350 of SEQ ID NO:7 is  
19 replaced with A; the K corresponding to position 371 of SEQ ID NO:2 or position 352 of  
20 SEQ ID NO:7 is replaced with A; the L corresponding to position 372 of SEQ ID NO:2  
21 or position 353 of SEQ ID NO:7 is replaced with A; and the L corresponding to position  
22 373 of SEQ ID NO:2 or position 354 of SEQ ID NO:7 is replaced with A.

23 More preferably, the recombinant host cell comprises a structural nucleic acid  
24 sequence encoding SEQ ID NO:2 modified by the following changes or SEQ ID NO:7  
25 modified by the following changes: the Y corresponding to position 106 of SEQ ID NO:2  
26 or position 87 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 129 of  
27 SEQ ID NO:2 or position 110 of SEQ ID NO:7 is replaced with F; the Y corresponding  
28 to position 185 of SEQ ID NO:2 or position 166 of SEQ ID NO:7 is replaced with F; the  
29 Y corresponding to position 193 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is  
30 replaced with F; the Y corresponding to position 270 of SEQ ID NO:2 or position 251 of  
31 SEQ ID NO:7 is replaced with F; the Y corresponding to position 316 of SEQ ID NO:2

1 or position 297 of SEQ ID NO:7 is replaced with F; and the Y corresponding to position  
2 362 of SEQ ID NO:2 or position 343 of SEQ ID NO:7 is replaced with F.

3 Most preferably, the recombinant host cell comprises a structural nucleic acid  
4 sequence encoding SEQ ID NO:2 modified by the following changes or SEQ ID NO:7  
5 modified by the following changes: the Y corresponding to position 185 of SEQ ID NO:2  
6 or position 166 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 193  
7 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is replaced with F; the Y  
8 corresponding to position 270 of SEQ ID NO:2 or position 251 of SEQ ID NO:7 is  
9 replaced with F; the Y corresponding to position 316 of SEQ ID NO:2 or position 297 of  
10 SEQ ID NO:7 is replaced with F; and the Y corresponding to position 362 of SEQ ID  
11 NO:2 or position 343 of SEQ ID NO:7 is replaced with F.

12 The recombinant host cell can generally be any type of host cell, and preferably is  
13 a bacterial, fungal, or plant cell. The bacterial cell is preferably an *Escherichia coli*  
14 bacterial cell. The fungal cell is preferably a *Saccharomyces cerevisiae*,  
15 *Schizosaccharomyces pombe*, or *Pichia pastoris* fungal cell. The plant cell can be a  
16 monocot, dicot, or conifer plant cell. The plant cell is preferably an alfalfa, banana,  
17 canola, corn, cotton, cucumber, peanut, potato, rice, soybean, sunflower, sweet potato,  
18 tobacco, tomato, or wheat plant cell. The recombinant host cell preferably further  
19 comprises operatively linked to the structural nucleic acid sequence a promoter that  
20 directs transcription of the structural nucleic acid sequence. The recombinant host cell  
21 preferably further comprises operatively linked to the structural nucleic acid sequence a  
22 3' transcription terminator and a polyadenylation site.

23 Recombinant plants

24 An additional embodiment of the invention is a recombinant plant comprising a  
25 structural nucleic acid sequence encoding a deallergenized patatin protein. The  
26 recombinant plant preferably produces a deallergenized patatin protein. More preferably,  
27 the recombinant plant produces a deallergenized patatin protein in a concentration  
28 sufficient to inhibit growth or to kill an insect which ingests plant tissue from the  
29 recombinant plant.

1           The recombinant plant can comprise a structural nucleic acid sequence encoding  
2 SEQ ID NO:2 modified in one or more of the following regions, or SEQ ID NO:7  
3 modified in one or more of the following regions. The single or multiple amino acid  
4 modifications reduce the binding of the modified protein relative to the binding of the  
5 corresponding unmodified protein. The regions for modification include amino acid  
6 positions 104-113 of SEQ ID NO:2 (85-94 of SEQ ID NO:7), 128-137 of SEQ ID NO:2  
7 (109-118 of SEQ ID NO:7), 184-197 of SEQ ID NO:2 (165-178 of SEQ ID NO:7), 264-  
8 277 of SEQ ID NO:2 (245-258 of SEQ ID NO:7), 316-325 of SEQ ID NO:2 (297-306 of  
9 SEQ ID NO:7), and 360-377 of SEQ ID NO:2 (341-358 of SEQ ID NO:7). The possible  
10 amino acid modifications include replacing an amino acid with A, E, F, P, or S. The  
11 modifications replace one or more amino acids in the identified regions, without  
12 increasing or decreasing the total number of amino acids in the protein.

13           Alternatively, the recombinant plant can comprise a structural nucleic acid  
14 sequence encoding SEQ ID NO:2 modified by one or more of the following changes or  
15 encoding SEQ ID NO:7 modified by one or more of the following changes: the Y  
16 corresponding to position 106 of SEQ ID NO:2 or position 87 of SEQ ID NO:7 is  
17 replaced with F or A; the I corresponding to position 113 of SEQ ID NO:2 or position 94  
18 of SEQ ID NO:7 is replaced with A; the Y corresponding to position 129 of SEQ ID  
19 NO:2 or position 110 of SEQ ID NO:7 is replaced with F or A; the K corresponding to  
20 position 137 of SEQ ID NO:2 or position 118 of SEQ ID NO:7 is replaced with A; the S  
21 corresponding to position 184 of SEQ ID NO:2 or position 165 of SEQ ID NO:7 is  
22 replaced with A; the Y corresponding to position 185 of SEQ ID NO:2 or position 166 of  
23 SEQ ID NO:7 is replaced with F or A; the A corresponding to position 188 of SEQ ID  
24 NO:2 or position 169 of SEQ ID NO:7 is replaced with S; the T corresponding to position  
25 192 of SEQ ID NO:2 or position 173 of SEQ ID NO:7 is replaced with A or P; the Y  
26 corresponding to position 193 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is  
27 replaced with F or A; the K corresponding to position 268 of SEQ ID NO:2 or position  
28 249 of SEQ ID NO:7 is replaced with A or E; the T corresponding to position 269 of  
29 SEQ ID NO:2 or position 250 of SEQ ID NO:7 is replaced with A; the Y corresponding  
30 to position 270 of SEQ ID NO:2 or position 251 of SEQ ID NO:7 is replaced with F or  
31 A; the K corresponding to position 273 of SEQ ID NO:2 or position 254 of SEQ ID NO:7

1 is replaced with A; the K corresponding to position 313 of SEQ ID NO:2 or position 294  
2 of SEQ ID NO:7 is replaced with E; the N corresponding to position 314 of SEQ ID NO:  
3 NO:2 or position 295 of SEQ ID NO:7 is replaced with A; the N corresponding to  
4 position 315 of SEQ ID NO:2 or position 296 of SEQ ID NO:7 is replaced with A; the Y  
5 corresponding to position 316 of SEQ ID NO:2 or position 297 of SEQ ID NO:7 is  
6 replaced with F or A; the Y corresponding to position 362 of SEQ ID NO:2 or position  
7 343 of SEQ ID NO:7 is replaced with F; the K corresponding to position 367 of SEQ ID  
8 NO:2 or position 348 of SEQ ID NO:7 is replaced with A; the R corresponding to  
9 position 368 of SEQ ID NO:2 or position 349 of SEQ ID NO:7 is replaced with A; the F  
10 corresponding to position 369 of SEQ ID NO:2 or position 350 of SEQ ID NO:7 is  
11 replaced with A; the K corresponding to position 371 of SEQ ID NO:2 or position 352 of  
12 SEQ ID NO:7 is replaced with A; the L corresponding to position 372 of SEQ ID NO:2  
13 or position 353 of SEQ ID NO:7 is replaced with A; and the L corresponding to position  
14 373 of SEQ ID NO:2 or position 354 of SEQ ID NO:7 is replaced with A.

15 More preferably, the recombinant plant comprises a structural nucleic acid  
16 sequence encoding SEQ ID NO:2 modified by the following changes or SEQ ID NO:7  
17 modified by the following changes: the Y corresponding to position 106 of SEQ ID NO:2  
18 or position 87 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 129 of  
19 SEQ ID NO:2 or position 110 of SEQ ID NO:7 is replaced with F; the Y corresponding  
20 to position 185 of SEQ ID NO:2 or position 166 of SEQ ID NO:7 is replaced with F; the  
21 Y corresponding to position 193 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is  
22 replaced with F; the Y corresponding to position 270 of SEQ ID NO:2 or position 251 of  
23 SEQ ID NO:7 is replaced with F; the Y corresponding to position 316 of SEQ ID NO:2  
24 or position 297 of SEQ ID NO:7 is replaced with F; and the Y corresponding to position  
25 362 of SEQ ID NO:2 or position 343 of SEQ ID NO:7 is replaced with F.

26 Most preferably, the recombinant plant comprises a structural nucleic acid  
27 sequence encoding SEQ ID NO:2 modified by the following changes or SEQ ID NO:7  
28 modified by the following changes: the Y corresponding to position 185 of SEQ ID NO:2  
29 or position 166 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 193  
30 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is replaced with F; the Y  
31 corresponding to position 270 of SEQ ID NO:2 or position 251 of SEQ ID NO:7 is

1 replaced with F; the Y corresponding to position 316 of SEQ ID NO:2 or position 297 of  
2 SEQ ID NO:7 is replaced with F; and the Y corresponding to position 362 of SEQ ID  
3 NO:2 or position 343 of SEQ ID NO:7 is replaced with F.

4 The recombinant plant can generally be any type of plant. The plant can be a  
5 monocot, dicot, or conifer plant. The plant is preferably an alfalfa, banana, canola, corn,  
6 cotton, cucumber, peanut, potato, rice, soybean, sunflower, sweet potato, tobacco,  
7 tomato, or wheat plant.

8 The recombinant plant preferably further comprises operatively linked to the  
9 structural nucleic acid sequence a promoter that directs transcription of the structural  
10 nucleic acid sequence. The recombinant plant preferably further comprises operatively  
11 linked to the structural nucleic acid sequence a 3' transcription terminator and a  
12 polyadenylation site.

13 Methods of preparation

14 Embodiments of the invention are further directed towards methods of preparing  
15 recombinant host cells and recombinant plants useful for the production of deallergenized  
16 patatin proteins.

17 A method of preparing a recombinant host cell useful for the production of  
18 deallergenized patatin proteins can comprise selecting a host cell; transforming the host  
19 cell with a recombinant vector; and obtaining recombinant host cells.

20 The recombinant vector comprises a structural nucleic acid sequence encoding  
21 SEQ ID NO:2 modified in one or more of the following regions, or SEQ ID NO:7  
22 modified in one or more of the following regions. The single or multiple amino acid  
23 modifications reduce the binding of the modified protein relative to the binding of the  
24 corresponding unmodified protein. The regions for modification include amino acid  
25 positions 104-113 of SEQ ID NO:2 (85-94 of SEQ ID NO:7), 128-137 of SEQ ID NO:2  
26 (109-118 of SEQ ID NO:7), 184-197 of SEQ ID NO:2 (165-178 of SEQ ID NO:7), 264-  
27 277 of SEQ ID NO:2 (245-258 of SEQ ID NO:7), 316-325 of SEQ ID NO:2 (297-306 of  
28 SEQ ID NO:7), and 360-377 of SEQ ID NO:2 (341-358 of SEQ ID NO:7). The possible  
29 amino acid modifications include replacing an amino acid with A, E, F, P, or S. The

1 modifications replace one or more amino acids in the identified regions, without  
2 increasing or decreasing the total number of amino acids in the protein.

3 Alternatively, the recombinant vector comprises a structural nucleic acid sequence  
4 encoding SEQ ID NO:2 modified by one or more of the following changes or encoding  
5 SEQ ID NO:7 modified by one or more of the following changes: the Y corresponding to  
6 position 106 of SEQ ID NO:2 or position 87 of SEQ ID NO:7 is replaced with F or A; the  
7 I corresponding to position 113 of SEQ ID NO:2 or position 94 of SEQ ID NO:7 is  
8 replaced with A; the Y corresponding to position 129 of SEQ ID NO:2 or position 110 of  
9 SEQ ID NO:7 is replaced with F or A; the K corresponding to position 137 of SEQ ID  
10 NO:2 or position 118 of SEQ ID NO:7 is replaced with A; the S corresponding to  
11 position 184 of SEQ ID NO:2 or position 165 of SEQ ID NO:7 is replaced with A; the Y  
12 corresponding to position 185 of SEQ ID NO:2 or position 166 of SEQ ID NO:7 is  
13 replaced with F or A; the A corresponding to position 188 of SEQ ID NO:2 or position  
14 169 of SEQ ID NO:7 is replaced with S; the T corresponding to position 192 of SEQ ID  
15 NO:2 or position 173 of SEQ ID NO:7 is replaced with A or P; the Y corresponding to  
16 position 193 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is replaced with F or A;  
17 the K corresponding to position 268 of SEQ ID NO:2 or position 249 of SEQ ID NO:7 is  
18 replaced with A or E; the T corresponding to position 269 of SEQ ID NO:2 or position  
19 250 of SEQ ID NO:7 is replaced with A; the Y corresponding to position 270 of SEQ ID  
20 NO:2 or position 251 of SEQ ID NO:7 is replaced with F or A; the K corresponding to  
21 position 273 of SEQ ID NO:2 or position 254 of SEQ ID NO:7 is replaced with A; the K  
22 corresponding to position 313 of SEQ ID NO:2 or position 294 of SEQ ID NO:7 is  
23 replaced with E; the N corresponding to position 314 of SEQ ID NO:2 or position 295 of  
24 SEQ ID NO:7 is replaced with A; the N corresponding to position 315 of SEQ ID NO:2  
25 or position 296 of SEQ ID NO:7 is replaced with A; the Y corresponding to position 316  
26 of SEQ ID NO:2 or position 297 of SEQ ID NO:7 is replaced with F or A; the Y  
27 corresponding to position 362 of SEQ ID NO:2 or position 343 of SEQ ID NO:7 is  
28 replaced with F; the K corresponding to position 367 of SEQ ID NO:2 or position 348 of  
29 SEQ ID NO:7 is replaced with A; the R corresponding to position 368 of SEQ ID NO:2  
30 or position 349 of SEQ ID NO:7 is replaced with A; the F corresponding to position 369  
31 of SEQ ID NO:2 or position 350 of SEQ ID NO:7 is replaced with A; the K

1 corresponding to position 371 of SEQ ID NO:2 or position 352 of SEQ ID NO:7 is  
2 replaced with A; the L corresponding to position 372 of SEQ ID NO:2 or position 353 of  
3 SEQ ID NO:7 is replaced with A; and the L corresponding to position 373 of SEQ ID  
4 NO:2 or position 354 of SEQ ID NO:7 is replaced with A.

5 More preferably, the vector comprises a structural nucleic acid sequence encoding  
6 SEQ ID NO:2 modified by the following changes or SEQ ID NO:7 modified by the  
7 following changes: the Y corresponding to position 106 of SEQ ID NO:2 or position 87  
8 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 129 of SEQ ID  
9 NO:2 or position 110 of SEQ ID NO:7 is replaced with F; the Y corresponding to  
10 position 185 of SEQ ID NO:2 or position 166 of SEQ ID NO:7 is replaced with F; the Y  
11 corresponding to position 193 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is  
12 replaced with F; the Y corresponding to position 270 of SEQ ID NO:2 or position 251 of  
13 SEQ ID NO:7 is replaced with F; the Y corresponding to position 316 of SEQ ID NO:2  
14 or position 297 of SEQ ID NO:7 is replaced with F; and the Y corresponding to position  
15 362 of SEQ ID NO:2 or position 343 of SEQ ID NO:7 is replaced with F.

16 Most preferably, the vector comprises a structural nucleic acid sequence encoding  
17 SEQ ID NO:2 modified by the following changes or SEQ ID NO:7 modified by the  
18 following changes: the Y corresponding to position 185 of SEQ ID NO:2 or position 166  
19 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 193 of SEQ ID  
20 NO:2 or position 174 of SEQ ID NO:7 is replaced with F; the Y corresponding to  
21 position 270 of SEQ ID NO:2 or position 251 of SEQ ID NO:7 is replaced with F; the Y  
22 corresponding to position 316 of SEQ ID NO:2 or position 297 of SEQ ID NO:7 is  
23 replaced with F; and the Y corresponding to position 362 of SEQ ID NO:2 or position  
24 343 of SEQ ID NO:7 is replaced with F.

25 The method can generally be used to prepare any type of recombinant host cell.  
26 Preferably, the method can be used to prepare a recombinant bacterial cell, a recombinant  
27 fungal cell, or a recombinant plant cell. The bacterial cell is preferably an *Escherichia*  
28 *coli* bacterial cell. The fungal cell is preferably a *Saccharomyces cerevisiae*,  
29 *Schizosaccharomyces pombe*, or *Pichia pastoris* fungal cell. The plant cell can be a  
30 monocot, dicot, or conifer plant cell. The plant cell is preferably an alfalfa, banana,

1 canola, corn, cotton, cucumber, peanut, potato, rice, soybean, sunflower, sweet potato,  
2 tobacco, tomato, or wheat plant cell.

3 An additional embodiment is directed towards methods for the preparation of  
4 recombinant plants useful for the production of deallergenized patatin proteins. The  
5 method can comprise selecting a host plant cell; transforming the host plant cell with a  
6 recombinant vector; obtaining recombinant host cells; and regenerating a recombinant  
7 plant from the recombinant host plant cells.

8 The recombinant vector comprises a structural nucleic acid sequence encoding  
9 SEQ ID NO:2 modified in one or more of the following regions, or SEQ ID NO:7  
10 modified in one or more of the following regions. The single or multiple amino acid  
11 modifications reduce the binding of the modified protein relative to the binding of the  
12 corresponding unmodified protein. The regions for modification include amino acid  
13 positions 104-113 of SEQ ID NO:2 (85-94 of SEQ ID NO:7), 128-137 of SEQ ID NO:2  
14 (109-118 of SEQ ID NO:7), 184-197 of SEQ ID NO:2 (165-178 of SEQ ID NO:7), 264-  
15 277 of SEQ ID NO:2 (245-258 of SEQ ID NO:7), 316-325 of SEQ ID NO:2 (297-306 of  
16 SEQ ID NO:7), and 360-377 of SEQ ID NO:2 (341-358 of SEQ ID NO:7). The possible  
17 amino acid modifications include replacing an amino acid with A, E, F, P, or S. The  
18 modifications replace one or more amino acids in the identified regions, without  
19 increasing or decreasing the total number of amino acids in the protein.

20 Alternatively, the recombinant vector comprises a structural nucleic acid sequence  
21 encoding SEQ ID NO:2 modified by one or more of the following changes or encoding  
22 SEQ ID NO:7 modified by one or more of the following changes: the Y corresponding to  
23 position 106 of SEQ ID NO:2 or position 87 of SEQ ID NO:7 is replaced with F or A; the  
24 I corresponding to position 113 of SEQ ID NO:2 or position 94 of SEQ ID NO:7 is  
25 replaced with A; the Y corresponding to position 129 of SEQ ID NO:2 or position 110 of  
26 SEQ ID NO:7 is replaced with F or A; the K corresponding to position 137 of SEQ ID  
27 NO:2 or position 118 of SEQ ID NO:7 is replaced with A; the S corresponding to  
28 position 184 of SEQ ID NO:2 or position 165 of SEQ ID NO:7 is replaced with A; the Y  
29 corresponding to position 185 of SEQ ID NO:2 or position 166 of SEQ ID NO:7 is  
30 replaced with F or A; the A corresponding to position 188 of SEQ ID NO:2 or position  
31 169 of SEQ ID NO:7 is replaced with S; the T corresponding to position 192 of SEQ ID

1 NO:2 or position 173 of SEQ ID NO:7 is replaced with A or P; the Y corresponding to  
2 position 193 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is replaced with F or A;  
3 the K corresponding to position 268 of SEQ ID NO:2 or position 249 of SEQ ID NO:7 is  
4 replaced with A or E; the T corresponding to position 269 of SEQ ID NO:2 or position  
5 250 of SEQ ID NO:7 is replaced with A; the Y corresponding to position 270 of SEQ ID  
6 NO:2 or position 251 of SEQ ID NO:7 is replaced with F or A; the K corresponding to  
7 position 273 of SEQ ID NO:2 or position 254 of SEQ ID NO:7 is replaced with A; the K  
8 corresponding to position 313 of SEQ ID NO:2 or position 294 of SEQ ID NO:7 is  
9 replaced with E; the N corresponding to position 314 of SEQ ID NO:2 or position 295 of  
10 SEQ ID NO:7 is replaced with A; the N corresponding to position 315 of SEQ ID NO:2  
11 or position 296 of SEQ ID NO:7 is replaced with A; the Y corresponding to position 316  
12 of SEQ ID NO:2 or position 297 of SEQ ID NO:7 is replaced with F or A; the Y  
13 corresponding to position 362 of SEQ ID NO:2 or position 343 of SEQ ID NO:7 is  
14 replaced with F; the K corresponding to position 367 of SEQ ID NO:2 or position 348 of  
15 SEQ ID NO:7 is replaced with A; the R corresponding to position 368 of SEQ ID NO:2  
16 or position 349 of SEQ ID NO:7 is replaced with A; the F corresponding to position 369  
17 of SEQ ID NO:2 or position 350 of SEQ ID NO:7 is replaced with A; the K  
18 corresponding to position 371 of SEQ ID NO:2 or position 352 of SEQ ID NO:7 is  
19 replaced with A; the L corresponding to position 372 of SEQ ID NO:2 or position 353 of  
20 SEQ ID NO:7 is replaced with A; and the L corresponding to position 373 of SEQ ID  
21 NO:2 or position 354 of SEQ ID NO:7 is replaced with A.

22 More preferably, the vector comprises a structural nucleic acid sequence encoding  
23 SEQ ID NO:2 modified by the following changes or SEQ ID NO:7 modified by the  
24 following changes: the Y corresponding to position 106 of SEQ ID NO:2 or position 87  
25 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 129 of SEQ ID  
26 NO:2 or position 110 of SEQ ID NO:7 is replaced with F; the Y corresponding to  
27 position 185 of SEQ ID NO:2 or position 166 of SEQ ID NO:7 is replaced with F; the Y  
28 corresponding to position 193 of SEQ ID NO:2 or position 174 of SEQ ID NO:7 is  
29 replaced with F; the Y corresponding to position 270 of SEQ ID NO:2 or position 251 of  
30 SEQ ID NO:7 is replaced with F; the Y corresponding to position 316 of SEQ ID NO:2

1 or position 297 of SEQ ID NO:7 is replaced with F; and the Y corresponding to position  
2 362 of SEQ ID NO:2 or position 343 of SEQ ID NO:7 is replaced with F.

3 Most preferably, the vector comprises a structural nucleic acid sequence encoding  
4 SEQ ID NO:2 modified by the following changes or SEQ ID NO:7 modified by the  
5 following changes: the Y corresponding to position 185 of SEQ ID NO:2 or position 166  
6 of SEQ ID NO:7 is replaced with F; the Y corresponding to position 193 of SEQ ID  
7 NO:2 or position 174 of SEQ ID NO:7 is replaced with F; the Y corresponding to  
8 position 270 of SEQ ID NO:2 or position 251 of SEQ ID NO:7 is replaced with F; the Y  
9 corresponding to position 316 of SEQ ID NO:2 or position 297 of SEQ ID NO:7 is  
10 replaced with F; and the Y corresponding to position 362 of SEQ ID NO:2 or position  
11 343 of SEQ ID NO:7 is replaced with F.

12 The recombinant plant can generally be any type of plant. The plant can be a  
13 monocot, dicot, or conifer plant. The plant is preferably an alfalfa, banana, canola, corn,  
14 cotton, cucumber, peanut, potato, rice, soybean, sunflower, sweet potato, tobacco,  
15 tomato, or wheat plant.

16 Deallergenized patatin proteins can be prepared by isolating the deallergenized  
17 patatin protein from any one of the above described host cells or plants.

18

#### 19 Deglycosylation

20 The examples herein provide evidence that glycosylation of can contribute to the  
21 allergenicity of a protein. Accordingly, rational substitution of amino acid residues likely  
22 to be the targets of glycosylation within a subject allergen protein may reduce or  
23 eliminate the allergenic properties of the protein without adversely affecting the  
24 enzymatic, insecticidal, antifungal or other functional properties of the protein.

25 Glycosylation commonly occurs as either N-linked or O-linked forms. N-linked  
26 glycosylation usually occurs at the motif Asn-Xaa-Ser/Thr, where Xaa is any amino acid  
27 except Pro (Kasturi, L. et al., *Biochem J.* 323: 415-519, 1997; Melquist, J.L. et al.,  
28 *Biochemistry* 37: 6833-6837, 1998). O-linked glycosylation occurs between the hydroxyl  
29 group of serine or threonine and an amino sugar.

30 Site directed mutagenesis of selected asparagine, serine, or threonine may be used  
31 to reduce or eliminate the glycosylation of patatin proteins. A search of SEQ ID NO:2

1 for the Asn-Xaa-Ser/Thr motif reveals one occurrence at amino acid positions 202-204.  
2 Mutagenization of the nucleic acid sequence encoding this region may result in a reduced  
3 allergenicity of the encoded protein.

4 In order to test this conceptual approach to reducing allergenicity of patatin  
5 proteins, two sets of experiments were performed: a) production of patatin proteins in  
6 *Escherichia coli*, which do not glycosylate proteins; and b) production of patatin proteins  
7 with an N202Q site directed mutation.

8 Antibodies obtained from patients HS-07 and G15-MON (not potato allergic) did  
9 not show specific binding to wild type patatin, patatin produced in *E. coli*, or the N202Q  
10 variant. Antibodies obtained from patient HS-01 (potato allergic) bound to wild type  
11 patatin, but not to patatin produced in *E. coli* or the N202Q variant. Antibodies obtained  
12 from patient HS-02 (potato allergic) bound strongly to wild type patatin, but extremely  
13 weakly to patatin produced in *E. coli*, and binding to the N202Q variant resembled vector  
14 controls. Antibodies obtained from patient HS-03 (potato allergic) bound to wild type  
15 patatin, but not to patatin produced in *E. coli* or the N202Q variant. Antibodies obtained  
16 from patient HS-05 (potato allergic) bound to wild type patatin, but very weakly to  
17 patatin produced in *E. coli* and the N202Q variant. Antibodies obtained from patient HS-  
18 06 (potato allergic) strongly bound wild type patatin, the N202Q variant, and to patatin  
19 produced in *E. coli*. These results strongly suggest that glycosylation is at least partially  
20 responsible for the antigenic properties of patatin proteins, and that site directed  
21 mutagenesis may be used to reduce or eliminate specific antibody binding. Mutagenesis  
22 at position 202 of SEQ ID NO:2 may be useful for reducing or eliminating specific  
23 antibody binding.

24 Permuteins

25 The positions of the internal breakpoints described in the following Examples are  
26 found on the protein surface, and are distributed throughout the linear sequence without  
27 any obvious bias towards the ends or the middle. Breakpoints occurring below the  
28 protein surface can additionally be selected. The rearranged two subunits can be joined  
29 by a peptide linker. A preferred embodiment involves the linking of the N-terminal and  
30 C-terminal subunits by a three amino acid linker, although linkers of various sizes can be

1 used. Additionally, the N-terminal and C-terminal subunits can be joined lacking a linker  
2 sequence. Furthermore, a portion of the C-terminal subunit can be deleted and the  
3 connection made from the truncated C-terminal subunit to the original N-terminal subunit  
4 and vice versa as previously described (Yang and Schachman, *Proc. Natl. Acad. Sci.*  
5 *U.S.A.*, 90: 11980-11984, 1993; Viguera, *et al.*, *Mol. Biol.*, 247: 670-681, 1995;  
6 Protasova, *et al.*, *Prot. Eng.*, 7: 1373-1377, 1994).

7 The novel insecticidal proteins of the present invention can be represented by the  
8 formula:

9  $X^1-(L)_a-X^2$

10 wherein;

11 a is 0 or 1, and if a is 0, then the permutein does not contain a linker  
12 sequence;

13  $X^1$  is a polypeptide sequence corresponding to amino acids n+1 through J;

14  $X^2$  is a polypeptide corresponding to amino acids 1 through n;

15 n is an integer ranging from 1 to J-1;

16 J is an integer greater than n+1; and

17 L is a linker.

18 In the formula above, the constituent amino acid residues of the novel insecticidal  
19 protein are numbered sequentially 1 through J from the original amino terminus to the  
20 original carboxyl terminus. A pair of adjacent amino acids within this protein can be  
21 numbered n and n+1 respectively where n is an integer ranging from 1 to J-1. The  
22 residue n+1 becomes the new N-terminus of the novel insecticidal protein and the residue  
23 n becomes the new C-terminus of the novel insecticidal protein.

24 For example, a parent protein sequence consisting of 120 amino acids can be  
25 selected as a starting point for designing a permutein (J=120). If the breakpoint is  
26 selected as being between position 40 and position 41, then n=40. If a linker is selected  
27 to join the two subunits, the resulting permutein will have the formula: (amino acids 41-  
28 120)-L-(amino acids 1-40). If a linker was not used, the resulting permutein will have the  
29 formula: (amino acids 41-120)-(amino acids 1-40).

30 The length of the amino acid sequence of the linker can be selected empirically,  
31 by using structural information, or by using a combination of the two approaches. When

1 no structural information is available, a small series of linkers can be made whose length  
2 can span a range of 0 to 50 Å and whose sequence is chosen in order to be substantially  
3 consistent with surface exposure (Hopp and Woods, *Mol. Immunol.*, 20: 483-489, 1983;  
4 Kyte and Doolittle, *J. Mol. Biol.*, 157: 105-132, 1982; Lee and Richards, *J. Mol. Biol.*,  
5 55: 379-400, 1971) and the ability to adopt a conformation which does not significantly  
6 affect the overall configuration of the protein (Karplus and Schulz, *Naturwissenschaften*,  
7 72: 212-213, 1985). Assuming an average length of 2.0 to 3.8 Å per residue, this would  
8 mean the length to test would be between about 0 to about 30 residues, with 0 to about 15  
9 residues being the preferred range. Accordingly, there are many such sequences that vary  
10 in length or composition that can serve as linkers with the primary consideration being  
11 that they be neither excessively long nor excessively short (Sandhu, *et al.*, *Critical Rev.*  
12 *Biotech.*, 12: 437-467, 1992). If the linker is too long, entropy effects may destabilize the  
13 three-dimensional fold and may affect protein folding. If the linker is too short, it may  
14 destabilize the molecule due to torsional or steric strain.

15 Use of the distance between the chain ends, defined as the distance between the  
16 C-alpha carbons, can be used to define the length of the sequence to be used, or at least to  
17 limit the number of possibilities that can be tested in an empirical selection of linkers.  
18 Using the calculated length as a guide, linkers with a range of number of residues  
19 (calculated using 2 to 3.8 Å per residue) can be selected. These linkers can be composed  
20 of the original sequence, shortened or lengthened as necessary, and when lengthened the  
21 additional residues can be chosen to be flexible and hydrophilic as described above; or  
22 optionally the original sequence can be substituted for using a series of linkers, one  
23 example being Gly-Pro-Gly (SEQ ID NO:277); or optionally a combination of the  
24 original sequence and new sequence having the appropriate total length can be used. An  
25 alternative short, flexible linker sequence is Gly-Gly-Gly-Ser-Gly-Gly-Gly (SEQ ID  
26 NO:276).

27 Selection of permutein breakpoints

28 Sequences of novel patatin analogs capable of folding to biologically active  
29 molecules can be prepared by appropriate selection of the beginning (amino terminus)  
30 and ending (carboxyl terminus) positions from within the original polypeptide chain

1 while optionally using a linker sequence as described above. Amino and carboxyl  
2 termini can be selected from within a common stretch of sequence, referred to as a  
3 breakpoint region, using the guidelines described below. A novel amino acid sequence is  
4 thus generated by selecting amino and carboxyl termini from within the same breakpoint  
5 region. In many cases, the selection of the new termini will be such that the original  
6 position of the carboxyl terminus immediately preceded that of the amino terminus.  
7 However, selections of termini anywhere within the region may result in a functional  
8 protein, and that these will effectively lead to either deletions or additions to the amino or  
9 carboxyl portions of the new sequence.

10 The primary amino acid sequence of a protein dictates folding to the three-  
11 dimensional structure beneficial for expression of its biological function. It is possible to  
12 obtain and interpret three-dimensional structural information using x-ray diffraction of  
13 single protein crystals or nuclear magnetic resonance spectroscopy of protein solutions.  
14 Examples of structural information that are relevant to the identification of breakpoint  
15 regions include the location and type of protein secondary structure (alpha and 3-10  
16 helices, parallel and anti-parallel beta sheets, chain reversals and turns, and loops  
17 (Kabsch and Sander, *Biopolymers*, 22: 2577-2637, 1983), the degree of solvent exposure  
18 of amino acid residues, the extent and type of interactions of residues with one another  
19 (Chothia, C., *Ann. Rev. Biochem.*, 53: 537-572, 1984), and the static and dynamic  
20 distribution of conformations along the polypeptide chain (Alber and Mathews, *Methods  
Enzymol.*, 154: 511-533, 1987). In some cases additional information is known about  
21 solvent exposure of residues, one example is a site of post-translational attachment of  
22 carbohydrate which is necessarily on the surface of the protein. When experimental  
23 structural information is not available, or when it is not feasible to obtain the information,  
24 methods are available to analyze the primary amino acid sequence in order to make  
25 predictions of protein secondary and tertiary structure, solvent accessibility and the  
26 occurrence of turns and loops (Fasman, G., Ed. Plenum, New York, 1989; Robson, B.  
27 and Garnier, J. *Nature* 361: 506, 1993).

29 Biochemical methods can be applicable for empirically determining surface  
30 exposure when direct structural methods are not feasible; for example, using the  
31 identification of sites of chain scission following limited proteolysis in order to infer

1 surface exposure (Gentile, F. and Salvatore, G., *Eur. J. Biochem.*, 218: 603-621, 1993).  
2 Thus, using either the experimentally derived structural information or predictive  
3 methods (Srinivasan, R. and Rose, G.D. *Proteins*, 22: 81-99, 1995), the parental amino  
4 acid sequence can be analyzed to classify regions according to whether or not they are  
5 integral to the maintenance of secondary and tertiary structure. The sequences within  
6 regions that are known to be involved in periodic secondary structure (alpha and 3-10  
7 helices, parallel and anti-parallel beta sheets) are regions that should be avoided.  
8 Similarly, regions of amino acid sequence that are observed or predicted to have a low  
9 degree of solvent exposure are more likely to be part of the so-called hydrophobic core of  
10 the protein and should also be avoided for selection of amino and carboxyl termini.  
11 Regions that are known or predicted to be in surface turns or loops, and especially those  
12 regions that are known not to be required for biological activity, can be preferred sites for  
13 new amino and carboxyl termini. Stretches of amino acid sequence that are preferred  
14 based on the above criteria can be selected as breakpoint regions.

15 An embodiment of the invention is directed towards patatin permutein proteins.  
16 The permutein proteins preferably maintain esterase activity and insecticidal properties.  
17 The permutein proteins preferably are less allergenic than the wild type patatin protein to  
18 individuals or animals allergic to potatoes. This can be assayed by the binding of  
19 antibodies to the wild type patatin and patatin permutein proteins.

20 The permutein proteins can optionally contain a linker sequence. The linker can  
21 generally be any amino acid sequence, preferably is Gly-Gly-Gly-Ser-Gly-Gly (SEQ  
22 ID NO:276) or Gly-Pro-Gly (SEQ ID NO:277), and more preferably is Gly-Pro-Gly  
23 (SEQ ID NO:277). Specific permutein proteins comprise: (amino acids 247-386 of SEQ  
24 ID NO:2)-linker-( amino acids 24-246 of SEQ ID NO:2), (amino acids 269-386 of SEQ  
25 ID NO:2)-linker-( amino acids 24-268 of SEQ ID NO:2), SEQ ID NO:247, and SEQ ID  
26 NO:259.

27 Embodiments of the invention also include isolated nucleic acid molecule  
28 segments comprising a structural nucleic acid sequence encoding a patatin permutein  
29 protein. The encoded permutein protein can generally be any permutein protein, and  
30 preferably comprises (amino acids 247-386 of SEQ ID NO:2)-linker-(amino acids 24-246  
31 of SEQ ID NO:2), (amino acids 269-386 of SEQ ID NO:2)-linker-(amino acids 24-268 of

1 SEQ ID NO:2), SEQ ID NO:247, or SEQ ID NO:259. The linker can generally be any  
2 amino acid sequence, preferably is Gly-Gly-Gly-Ser-Gly-Gly-Gly (SEQ ID NO:276) or  
3 Gly-Pro-Gly (SEQ ID NO:277), and more preferably is Gly-Pro-Gly (SEQ ID NO:277).  
4 Alternatively, the encoded patatin permutein protein can lack a linker sequence. The  
5 structural nucleic acid sequence is preferably SEQ ID NO:246 or SEQ ID NO:258.

6 An embodiment of the invention is directed towards recombinant vectors which  
7 encode a patatin permutein protein. The vector can comprise operatively linked in the 5'  
8 to 3' orientation: a promoter that directs transcription of a structural nucleic acid  
9 sequence; a structural nucleic acid sequence encoding a protein selected from the group  
10 consisting of: (amino acids 247-386 of SEQ ID NO:2)-linker-(amino acids 24-246 of  
11 SEQ ID NO:2); and (amino acids 269-386 of SEQ ID NO:2)-linker-(amino acids 24-268  
12 of SEQ ID NO:2); and a 3' transcription terminator. The linker can comprise Gly-Pro-  
13 Gly (SEQ ID NO:277) or Gly-Gly-Gly-Ser-Gly-Gly-Gly (SEQ ID NO:276).  
14 Alternatively, the encoded patatin permutein protein can lack a linker sequence. The  
15 structural nucleic acid sequence can preferably be SEQ ID NO:246 or SEQ ID NO:258,  
16 and preferably encodes SEQ ID NO:247 or SEQ ID NO:259.

17 An additional embodiment of the invention is directed towards recombinant host  
18 cells useful for the production of a patatin permutein protein. The recombinant host cell  
19 preferably produces a patatin permutein protein. More preferably, the recombinant host  
20 cell produces a patatin permutein protein in a concentration sufficient to inhibit growth or  
21 to kill an insect which ingests the recombinant host cell. The recombinant host cell can  
22 comprise a structural nucleic acid sequence encoding a protein selected from the group  
23 consisting of: (amino acids 247-386 of SEQ ID NO:2)-linker-(amino acids 24-246 of  
24 SEQ ID NO:2); and (amino acids 269-386 of SEQ ID NO:2)-linker-(amino acids 24-268  
25 of SEQ ID NO:2). The linker can generally be any amino acid sequence, and preferably  
26 is Gly-Pro-Gly (SEQ ID NO:277) or Gly-Gly-Gly-Ser-Gly-Gly-Gly (SEQ ID NO:276).  
27 Alternatively, the encoded patatin permutein protein can lack a linker sequence. The  
28 structural nucleic acid sequence is preferably SEQ ID NO:246 or SEQ ID NO:258, and  
29 preferably encodes SEQ ID NO:247 or SEQ ID NO:259. The structural nucleic acid  
30 sequence can be operatively linked to a promoter sequence that directs transcription of  
31 the structural nucleic acid sequence, a 3' transcription terminator, and a 3'

1 polyadenylation signal sequence. The recombinant host cell can generally be any type of  
2 host cell, and preferably is a bacterial, fungal, or plant host cell. The bacterial cell is  
3 preferably an *Escherichia coli* bacterial cell. The fungal cell is preferably a  
4 *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, or *Pichia pastoris* fungal cell.  
5 The plant cell can be a monocot, dicot, or conifer plant cell. The plant cell is preferably  
6 an alfalfa, banana, canola, corn, cotton, cucumber, peanut, potato, rice, soybean,  
7 sunflower, sweet potato, tobacco, tomato, or wheat plant cell.

8 An additional embodiment of the invention is directed towards recombinant plants  
9 which are useful for the production of patatin permutein proteins. The recombinant plant  
10 preferably produces a patatin permutein protein. More preferably, the recombinant plant  
11 produces a patatin permutein protein in a concentration sufficient to inhibit growth or to  
12 kill an insect which ingests tissue from the recombinant plant. The recombinant plant can  
13 comprise a structural nucleic acid sequence encoding a protein selected from the group  
14 consisting of: (amino acids 247-386 of SEQ ID NO:2)-linker-(amino acids 24-246 of  
15 SEQ ID NO:2); and (amino acids 269-386 of SEQ ID NO:2)-linker-(amino acids 24-268  
16 of SEQ ID NO:2). The linker can comprise Gly-Pro-Gly (SEQ ID NO:277) or Gly-Gly-  
17 Gly-Ser-Gly-Gly (SEQ ID NO:276). Alternatively, the encoded protein can lack a  
18 linker sequence. The structural nucleic acid sequence is preferably SEQ ID NO:246 or  
19 SEQ ID NO:258, and preferably encodes SEQ ID NO:247 or SEQ ID NO:259. The  
20 structural nucleic acid sequence can be operatively linked to a promoter sequence that  
21 directs transcription of the structural nucleic acid sequence, a 3' transcription terminator,  
22 and a 3' polyadenylation signal sequence. The recombinant plant can generally be any  
23 type of plant, and preferably is an alfalfa, banana, canola, corn, cotton, cucumber, peanut,  
24 potato, rice, soybean, sunflower, sweet potato, tobacco, tomato, or wheat plant.

25 Permutein proteins can be prepared by isolating the permutein protein from any  
26 one of the above described host cells or plants.

27 Immunotherapy for potato allergy

28 Immunotherapy for food allergy has been largely unsuccessful due to the lack of  
29 appropriate therapeutic reagents (Sampson, H.A., *J. Allergy Clin. Immunol.*, 90(2): 151-  
30 152, 1992). Immunotherapy has typically involved the administration (orally or by

1 subcutaneous injections) of increasing doses of crude protein extracts of the offending  
2 allergenic entities which usually contain variable mixes of many different proteins  
3 (Scheiner, O., *Wien Klin Wochenschr.*, 105(22): 653-658, 1993). While there are reports  
4 of highly successful clinical applications of immunotherapy for food allergens (Romano,  
5 P.C., *et al.*, *Allergol. Immunopathol. (Madr.)*, 12(4): 275-281, 1984), those reports are rare  
6 and the clinical literature in general recommends avoidance far more strongly than  
7 therapy (Gay, G., *Allerg. Immunol. (Paris)*, 29(6): 169-170, 1997). One of the primary  
8 reasons for the failure of many clinical attempts to induce tolerance to allergens in  
9 general and food allergens in particular relates to anecdotal comments by numerous  
10 allergists, that patients don't tolerate the doses of allergen required to achieve tolerance.  
11 Animal studies examining the relationship of antigen dose and the induction of tolerance  
12 have demonstrated a strong positive correlation (Chen, Y., *et al.*, *Proc. Natl. Acad. Sci.*,  
13 *U.S.A.*, 93: 388-391, 1996; Tokai, T., *et al.*, *Nat. Biotechnol.*, 15(8): 754-758, 1997). Due  
14 to the very real possibility of inducing an anaphylactic reaction in patients with native  
15 allergen, most clinical therapists are quite hesitant to use high doses therapeutically and  
16 are therefore compromising the likelihood of successful therapy.

17 In recent reports, recombinant technology has been used to reduce the allergenic  
18 potential of a major allergen without modifying the T cell epitopes, and allowing higher  
19 doses of protein to be used in therapy (Tokai, T., *et al.*, *Nat. Biotechnol.*, 15(8): 754-758,  
20 1997). In addition, a lack of understanding about the appropriate route of administration,  
21 the uncertainty of mechanisms responsible for induction of allergy and the uncertainty of  
22 mechanisms by which immunotherapy suppresses or blocks the T cell-IgE-Eosinophil/mast cell cycle have contributed to the large number of equivocal studies and  
23 clinical trials. Recent studies in animal models dealing with mechanisms, routes of  
24 administration, adjuvants and vaccine formulations have increased the likelihood that  
25 immunotherapy for allergies, including food allergies, will become a reproducibly  
26 successful clinical treatment when the appropriate therapeutic reagents are available  
27 (Sampson, H.A. and Burks, A.W., *Annu. Rev. Nutr.*, 16: 161-177, 1996; Kaminogawa, S.,  
28 *Biosci. Biotechnol. Biochem.*, 60(11): 1749-1756, 1996; Chapman, M.D., *et al.*, *Allergy*,  
29 52: 374-379, 1997; Barbeau, W.E., *Adv. Exp. Med. Biol.*, 415: 183-193, 1997; Cao, Y., *et*  
30 *al.*, *Immunology*, 90(1): 46-51, 1997; Garside, P. and Mowat, A.M., *Crit. Rev. Immunol.*,

1 17(2): 119-137, 1997; Rothe, M.J. and Grant-Kels, J.M., *J. Am. Acad. Dermatol.*, 35(1):  
2 1-13, 1996; Strobel, S., *Allergy*, 50(20): 18-25, 1995; Kruisbeek, A.M. and Amsen, D.,  
3 *Curr. Opin. Immunol.*, 8(2): 233-244, 1996; Herz, U., *et al.*, *Adv. Exp. Med. Biol.*, 409:  
4 25-32, 1996; Litwin, A., *et al.*, *J. Allergy Clin. Immunol.*, 100: 30-38, 1997;  
5 Vandewalker, M.L., *Mo. Med.*, 94(7): 311, 1997; Marshall, G.D., Jr. and Davis, F., *Nat.*  
6 *Biotechnol.*, 15(8): 718-719, 1997; Van Deusen, M.A., *et al.*, *Ann. Allergy Asthma*  
7 *Immunol.*, 78: 573-580, 1997; Jacobsen, L., *et al.*, *Allergy*, 52: 914-920, 1997, Scheiner,  
8 O. and Kraft, D., *Allergy* 50(5): 384-391, 1995.

9 Relative to immunotherapy, the critical aspects of the modified patatin genes  
10 described in this patent are that they can be used to synthesize purified, deallergenized-  
11 protein which can be used for patatin (potato) specific immunotherapy, with reduced  
12 potential for adverse and potentially fatal anaphylactic reactions in human or veterinary  
13 patients who have allergies to patatin or potatoes. Various strategies, including fixing or  
14 cross linking allergens, encapsulation of allergen for oral delivery, the use of small, T-cell  
15 epitope peptides and most recently, the use of engineered recombinant proteins, or  
16 modified gene vaccines are being tested in attempts to decrease the potential for  
17 anaphylactic reactions while inducing tolerance (Cao, Y., *et al.*, *Immunology*, 90(1): 46-  
18 51, 1997; Chapman, M.D., *et al.*, *Allergy*, 52: 374-379, 1997; Chapman, M.D., *et al.*, *Int.*  
19 *Arch. Allergy Immunol.*, 113(1-3): 102-104, 1997; Collins, S.P., *et al.*, *Clin. Exp. Allergy*,  
20 26(1): 36-42, 1996; Takai, T., *et al.*, *Mol. Immunol.*, 34(3): 255-261, 1997; Takai, T., *et*  
21 *al.*, *Nat. Biotechnol.*, 15(8) 754-758, 1997; Jirapongsananruk, O. and Leung, D.Y.M.,  
22 *Ann. Allergy Asthma Immunol.*, 79: 5-20, 1997; Litwin, A., *et al.*, *J. Allergy Clin.*  
23 *Immunol.*, 100: 30-38, 1997; Vandewalker, M.L., *Mo. Med.*, 94(7): 311, 1997; Raz, E., *et*  
24 *al.*, *Proc. Natl. Acad. Sci., U.S.A.*, 93: 5141-5145, 1996; Hoyne, G.F., *et al.*, *Clin.*  
25 *Immunol. Immunopathol.*, 80: S23-30, 1996; Hoyne, G.F., *et al.*, *Int. Immunol.*, 9(8):  
26 1165-1173, 1997; Vrtala, S., *et al.*, *J. Clin. Invest.*, 99(7): 1673-1681, 1997; Sato, Y., *et*  
27 *al.*, *Science*, 273: 352-354, 1996; Lee, D.J., *et al.*, *Int. Arch. Allergy Immunol.*, 113(1-3):  
28 227-230, 1997; Tsitoura, D.C., *et al.*, *J. Immunol.*, 157(5): 2160-2165, 1996; Hsu, C.H.,  
29 *et al.*, *Int. Immunol.*, 8(9):1405-1411, 1996; Hsu, C.H., *et al.*, *Nat. Med.*, 2(5): 540-544,  
30 1996).

1       The instant invention uses an engineered patatin protein, as expressed in any  
2 living cell, with or without post-synthesis modifications, for immunotherapy by the  
3 routes of cutaneous or subcutaneous exposure, injection, or by oral, gastro-intestinal,  
4 respiratory or nasal application, either with, or without the use of specific carriers,  
5 vehicles and adjuvants. The direct application of nucleic acid encoding recombinant  
6 patatin as the *in vivo* (in the patient) expression template (gene) as RNA-, DNA- or gene-  
7 vaccines is also the intended use of the engineered genetic materials defined here, coding  
8 for patatin, but with modified IgE binding sites. It is also the intent of this patent to cover  
9 the use of these modified genes described here including insertion into various DNA  
10 vectors including adenovirus, retrovirus, pox virus and replicating or non-replicating  
11 eukaryotic expression plasmids (Lee, D.J., *et al.*, *Int. Arch. Allergy Immunol.*, 113(1-3):  
12 227-230, 1997) with various promoters and regulatory sequences, which can be inserted  
13 into the patient's somatic cells (dendritic cells, epithelial cells, muscle fiber-cells,  
14 fibroblasts, etc.) for the purpose of expressing the recombinant gene product to alter the  
15 patient's immune response to the patatin proteins (Lee D.J., *et al.*, *Int. Arch. Allergy*  
16 *Immunol.*, 113(1-3): 227-230, 1997). Potential routes of administration foreseen in this  
17 application include previously described methods of encapsulation, emulsion, receptor or  
18 membrane fusion mediated uptake and methods of direct permeabilization or insertion of  
19 the DNA or corresponding RNA into the host cells.

20       The following examples are included to demonstrate preferred embodiments of  
21 the invention. It should be appreciated by those of skill in the art that the techniques  
22 disclosed in the examples which follow represent techniques discovered by the inventors  
23 to function well in the practice of the invention, and thus can be considered to constitute  
24 preferred modes for its practice. However, those of skill in the art should, in light of the  
25 present disclosure, appreciate that many changes can be made in the specific  
26 embodiments which are disclosed and still obtain a like or similar result without  
27 departing from the spirit and scope of the invention.

## EXAMPLES

## 2 Example 1: Identification of patatin as an allergen

3 Since patatin is commonly obtained from an allergenic source (potato), it was  
4 hypothesized that patatins in fact encode an important class of offending potato allergens  
5 (patatin was reported as allergenic by Seppala, U. et al., *J. Allergy Clin. Immunol.* 103:  
6 165-171, 1999). Assessment of potential allergens preferably include appropriate *in vitro*  
7 testing for IgE binding, in this case with potato allergic sera (Fuchs, R.L. and Astwood,  
8 J.D., *Food Technology*, 50: 83-88, 1996; Astwood, J.D., et al., *Monographs in allergy*  
9 Vol. 32: *Highlights in food allergy*, pp. 105-120, 1996, Metcalfe, D.D., et al., *Critical*  
10 *Reviews in Food Science and Nutrition*, 36S: 165-186, 1996). It is the recommendation  
11 of a working group organized by the IFBC and the ILSI Allergy and Immunology  
12 Institute that proteins encoded by nucleic acid sequences from allergenic sources such as  
13 potato (a "less-commonly" allergenic source) should be examined for their ability to react  
14 with IgEs of potato-allergic patients using a minimum of five individual patient sera  
15 (Metcalfe, D.D., et al., *Critical Reviews in Food Science and Nutrition*, 36S: 165-186,  
16 1996). Patatin-17 protein was tested for IgE binding using standard *in vitro* testing with  
17 serum taken from patients with bona fide well defined clinically displayed potato allergy  
18 as described below.

## 19 Clinical Characterization of Potato Allergic Subjects (Serum donors)

20 Patients who suffer from potato allergy were identified at Johns Hopkins Clinic  
21 (Baltimore, MD) and were evaluated for potato allergy using clinical criteria outlined in  
22 Table 2.

Serum was obtained from patients with convincing clinical history of potato allergy. The convincing history was defined as being one or more of the following: a) positive potato allergic as evaluated by double-blind placebo-control food challenge b) anaphylaxis and/or hospitalization due to the consumption of potatoes or c) dramatic skin test results.

1 Table 2: Clinical patient data

Patient	Clinical History	Flare/Wheal (Skin prick test)	DBPCFC (potato)
HS01	Most recent hospitalization: 10/19/93 AD, A, AR, FH, MFS, IgE = 1397 KIAUa/L	7/19, 4/14, 7/17	Not performed
HS02	Most recent hospitalization: 6/94 AD, FH, Latex (+) RAST, MFS, IgE = 7544K/L	20/26	Not performed
HS03	Most recent hospitalization: 7/27/95 AD, A, FH, MFS, IgE = N/A	5/13	Yes
HS05	Most recent hospitalization 5/30/95 AD, A, FH, MFS, IgE = 12341 ng/ml	4/9	Yes
HS06	Most recent hospitalization 6/13/95 AD, A FH, MFS IgE = N/A	5/20, 4/13, 5/12	Yes
HS07	Not potato allergic, allergic to egg, milk, peanuts, seafood. AD, A, AR, FH, MFS	High IgE control serum, not allergic to potato.	
HS08	Non-atopic (normal)	Low IgE control serum	

2

3 AD= Atopic dermatitis; FH= Food hypersensitivity; AR = Allergic rhinitis; A=

4 Asthma; MFS= Multiple food sensitivity; N/A = not available.

5 Example 2: Western blotting of patatin proteins

6 Western blotting experiments were performed using patatin protein purified to  
 7 near homogeneity from corn plants genetically engineered to produce patatin, patatin  
 8 producing crude genetically engineered corn leaf extracts, crude potato tuber extracts,  
 9 and non-transgenic corn leaf samples.

10 Protein samples were electrophoresed by SDS-PAGE (Laemmli, U.K., *Nature*  
 11 227: 680-685, 1970) and were electroblotted onto nitrocellulose. Protein blots were  
 12 processed by standard Western blotting (immunoblotting) techniques and were incubated  
 13 in potato allergic serum diluted 1:5 in PBS buffer for 1 hour. After washing the blots 3  
 14 times with PBS, the blots were incubated in biotinylated anti-IgE (Johns Hopkins  
 15 Hospital, Baltimore, MD) for 1 hour, followed by a 30 minute incubation in HRP-linked  
 16 avidin (Promega, New York, NY). IgE-reactive protein bands were visualized by DAB  
 17 staining (3,3 diaminobenzidine). The blots were dried and photographed. Individual

1 blots are labeled according to patient serum used. As a control, one blot was incubated in  
2 anti-IgE only.

3 Patatins were shown to be an allergen of potato by examining the reactivity of  
4 purified patatin to sera obtained from patients allergic to potato. Sera from five potato  
5 allergic subjects were tested by Western blotting techniques. All five sera reacted with  
6 purified patatin protein.

7 Patatin isozymes (SEQ ID NOS:278-282, Figure 1) were tested for IgE binding  
8 by Western blotting. Isozymes of patatin were cloned into a yeast expression system and  
9 purified prior to analysis. The isozymes were subjected to IgE western blotting as  
10 described above with the exception that all five patient sera were pooled. The resulting  
11 Western blot of the yeast-expressed isozymes showed that all five isozymes bound IgE in  
12 a manner similar to patatin 17, and that all isozymes of patatin tested are also allergens.

13 Example 3: Western blotting of patatin proteins

14 Eighty-nine 10-mer peptides were synthesized using the Genosys SPOTs system,  
15 each consecutive 10-mer overlapping by 6 amino acids based on the amino acid sequence  
16 of patatin 17 (SEQ ID NO:2). The peptides were evaluated for IgE binding with five  
17 different potato allergic patient sera using the same incubation procedures as described  
18 above. The results are summarized graphically in Figure 2, showing major and minor  
19 allergenic epitopes. Interestingly, many of the immunogenic epitopes contain tyrosine.  
20 The peptide numbers, sequences, and immunoreactivity is detailed in Table 3.

21 Table 3: Peptide scan of patatin 17

Peptide # (SEQ ID NO)	Peptide Sequence	HS01	HS02	HS03	HS05	HS06	Cumulative Total
1 (16)	QLGEMVTVLS	0.47	0.33	0.02	0.05	0.06	0.93
2 (17)	MVTVLSIDGG	0.53	0.33	0.02	0.07	0.05	1
3 (18)	LSIDGGGGIRG	0.52	0.38	0.07	0.08	0.09	1.14
4 (19)	GGGIRGIIPA	0.53	0.19	0.06	0.19	0.23	1.2
5 (20)	RGIPIATILE	0.46	0.28	0.04	0.09	0.05	0.92
6 (21)	PATILEFLEG	0.49	0.31	0.05	0.09	0.07	1.01
7 (22)	LEFLEGQLQE	0.36	0.24	0.04	0.1	0.06	0.8
8 (23)	EGQLQEMDNN	0.29	0.19	0.02	0.09	0.05	0.64
9 (24)	QEMDNNAADAR	0.22	0.13	0.01	0.05	0.04	0.45
10 (25)	NNADARLADY	0.21	0.17	0.03	0.05	0.07	0.53

11 (26)	ARLADYFDVI	0.54	0.31	0.16	0.15	0.25	1.41
12 (27)	DYFDVIGGTS	0.61	0.34	0.46	0.06	0.15	1.62
13 (28)	VIGGTSTGGL	0.63	0.72	0.05	0.15	0.09	1.64
14 (29)	TSTGGLLTAM	0.3	0.17	0.03	0.06	0.09	0.65
15 (30)	GLLTAMISTP	0.63	0.41	0.05	0.24	0.12	1.45
16 (31)	AMISTPNENN	0.34	0.18	0.02	0.07	0.02	0.63
17 (32)	TPNENNRPFA	0.46	0.22	0.03	0.19	0.07	0.97
18 (33)	NNRPFAAAKE	0.37	0.21	0.05	0.07	0.06	0.76
19 (34)	FAAAKEIVPF	0.52	0.29	0.08	0.11	0.08	1.08
20 (35)	KEIVPFYFEH	0.29	0.14	0.28	0.29	0.23	1.23
21 (36)	PFYFEGHPQI	0.65	0.06	1.08	0.51	0.17	2.47
22 (37)	EHGPOIFNPS	0.34	0.15	0.03	0.05	0.06	0.63
23 (38)	QIFNPSGQIL	0.33	0.29	0.02	0.07	0.07	0.78
24 (39)	PSGQILGPKY	0	0	0.02	0	0.05	0.07
25 (40)	ILGPKYDGGKY	0	0	0.07	0	0.02	0.09
26 (41)	KYDGKYLQMVG	0.02	0	0.11	0.01	0.04	0.18
27 (42)	KYLMQVQLQEK	0.12	0.04	1.08	0.07	0.79	2.1
28 (43)	QVQLQEKLGET	0.46	0.16	0.01	0.07	0.02	0.72
29 (44)	EKLGETRVRHQ	0.5	0.12	0.01	0.07	0.04	0.74
30 (45)	ETRVHQALTE	0.42	0.16	0.03	0.05	0.03	0.69
31 (46)	HQALTEVVVIS	0.43	0.21	0.04	0.1	0.05	0.83
32 (47)	TEVVISSFFDI	0.44	0.25	0.05	0.08	0.04	0.86
33 (48)	ISSFDIKTNK	0.1	0.02	0.04	0.06	0.13	0.35
34 (49)	DIKTNKPVIF	0.57	0.22	0.04	0.18	0.28	1.29
35 (50)	NKPVIFTKSN	0	0.01	0.02	0.07	0.24	0.34
36 (51)	IFTKSNLANS	0	0	0.03	0.06	0.17	0.26
37 (52)	SNLANSPELD	0.43	0.96	0.01	0.09	0.02	1.51
38 (53)	NSPELDIADKY	0.18	0.12	0.01	0.05	0.05	0.41
39 (54)	LDAKMYDISY	0.54	0.26	0.19	0.15	0.23	1.37
40 (55)	MYDISYSTAA	0.92	0.08	0.52	0.04	0.22	1.78
41 (56)	SYSTAAAPTY	1.15	0.25	1.04	0.33	0.55	3.32
42 (57)	AAAPTYFPFH	1.02	0.52	1.12	0.81	0.86	4.33
43 (58)	TYFPHYFVFT	0.02	0.01	0.54	0.03	0.24	0.84
44 (59)	PHYFVTNTSN	0.03	0.01	1.17	0.13	0.44	1.78
45 (60)	VTNTNSGDEY	0.23	0.15	0.04	0.03	0.03	0.48
46 (61)	SNGDEYEFLN	0.33	0.25	0.08	0.1	0.11	0.87
47 (62)	EYEFLNLDGA	0.34	0.25	0.07	0.1	0.2	0.96
48 (63)	NLVDGAVATV	0.3	0.18	0.02	0.06	0.05	0.61
49 (64)	GAVATVADPA	0.45	0.54	0.01	0.07	0.02	1.09
50 (65)	TVADPALLSI	0.48	0.29	0.01	0.07	0.03	0.88
51 (66)	PALLSISVAT	0.65	0.33	0.02	0.1	0.01	1.11
52 (67)	SISVATRLAQ	0.61	0.23	0.14	0.53	0.53	2.04
53 (68)	ATRLAQKDPA	0.87	0.34	0.05	0.29	0.22	1.77
54 (69)	AQKDPAFASI	0.86	0.32	0.04	0.12	0.03	1.37
55 (70)	PAFAISISSLN	0.81	0.15	0.05	0.51	0.59	2.11
56 (71)	SIRSLNYKKM	0.07	0.01	0.17	0.07	0.11	0.43
57 (72)	LNYKKMLLLS	0.05	0.01	0.35	0.08	0.39	0.88
58 (73)	KMLLLSLGTG	1.15	0.15	0.04	0.38	0.71	2.43
59 (74)	LSLGTGTTSE	0.34	0.23	0.02	0.04	0.03	0.66
60 (75)	TGTSEFDKT	0.92	0.39	0.6	0.1	0.09	2.1
61 (76)	SEFDKTYTAK	1.33	1.35	1.41	0.12	0.28	4.49
62 (77)	KTYTAKEAAT	1.36	0.94	1.11	0.76	0.4	4.57

63 (78)	AKEAATWTA	0.45	0.15	0.01	0.2	0.03	0.88
64 (79)	ATWTAVHWML	0.1	0.02	0.01	0.08	0.03	0.27
65 (80)	AVHWMVLVIQK	0.69	0.05	0.03	0.43	0.62	1.82
66 (81)	MLVIQKMTDA	0.32	0.15	0.02	0.15	0.03	0.67
67 (82)	QKMTDYYLST	0.26	0.125	0.03	0.21	0.03	0.675
68 (83)	DAASSYMTDY	0.2	0.14	0.08	0.08	0.1	0.6
69 (84)	SYMTDYYLST	0.5	0.03	0.32	0.05	0.11	1.02
70 (85)	DYYLSTAFQAA	0.44	0	0.22	0.43	0.03	0.52
71 (86)	STAFQALDSK	0.1	0.3	0.08	0.08	0.03	0.88
72 (87)	QALDSKNNYL	0.44	0.46	0.28	0.26	0.03	1.87
73 (88)	SKNNYLRVQE	0.44	0.05	1.31	0.07	0.21	2.08
74 (89)	YLRVQENALT	1.38	0.03	1.31	0.11	0.2	3.03
75 (90)	QENALTGTTT	0.47	0.25	0	0.06	0	0.78
76 (91)	LTGTITTEMDD	0.41	0.24	0	0.06	0	0.71
77 (92)	TTEMDDASEA	0.38	0.3	0	0.05	0	0.73
78 (93)	DDASEANMEL	0.44	0.24	0	0.06	0	0.74
79 (94)	EANMELLVQV	0.42	0.27	0	0.04	0	0.73
80 (95)	ELLVQVGENL	0.4	0.25	0	0.05	0	0.7
81 (96)	QVGENLKKP	0.44	0.14	0	0.07	0	0.65
82 (97)	NLLKKPVSED	0.47	0.2	0	0.03	0	0.7
83 (98)	KPVSEDNPET	0.27	0.21	0	0.03	0	0.51
84 (99)	EDNPETYEEA	0.13	0.11	0	0.01	0	0.25
85 (100)	ETYEAKLKR	1.26	1.2	1.36	0.53	0.71	5.06
86 (101)	EALKRFAKLL	1.38	0.04	0	1.06	1.12	3.6
87 (102)	RFAKLLSDRK	0.98	0.05	0	0.84	0.94	2.81
88 (103)	LLSDRKKLRA	0.2	0.01	0	0.37	0.51	1.09
89 (104)	RKKLRANKAS	0.28	0	0	0.31	0.64	1.23
	Patient Cumulative Totals	41.84	20.565	18.1	14.17	16.55	

1

2     Example 4: Identification of result effective substitutions

3     For each major and several minor allergenic epitopes of patatin, result effective  
 4     substitutions were identified by synthesizing peptides that were altered by individually  
 5     substituting an alanine residue at each non-alanine position in the epitope. Similarly, the  
 6     reported nucleic acid sequence encoding corn patatin (U.S. Patent No. 5,882,668; clone  
 7     5c9) was evaluated for IgE binding by producing peptides at corresponding positions to  
 8     the potato patatin protein.

9     For example, Epitope 41 was analyzed by alanine scanning and rational  
 10    substitution as follows.

11            Epitope 41                           SEFDKTYTAK (SEQ ID NO:76)  
 12            Alanine scan                       AEFDKTYTAK (SEQ ID NO:165)

1 SAFDKTYTAK (SEQ ID NO:166)  
 2 SEADKTYTAK (SEQ ID NO:167)  
 3 SEFAKTYTAK (SEQ ID NO:168)  
 4 SEFDATYTAK (SEQ ID NO:169)  
 5 SEFDKAYTAK (SEQ ID NO:170)  
 6 SEFDKTATAK (SEQ ID NO:171)  
 7 SEFDKTYAAK (SEQ ID NO:172)  
 8 SEFDKTYTAA (SEQ ID NO:173)  
 9 Rational substitution AFFDKTYTAK (SEQ ID NO:283)  
 10 SEFDKTFTAK (SEQ ID NO:176)  
 11 Corn homolog CIFDSTYTAK (SEQ ID NO:284)  
 12

13 Selected epitopes were analyzed by alanine scanning and rational substitution.  
 14 Immunoassay with potato-allergic serum was used as described above. Table 4  
 15 summarizes the results of these experiments to identify result effective substitutions for  
 16 patatin. Blank spaces in the table indicate that binding of the peptide to patient IgE was  
 17 not detectable.

18 Table 4: Scanning of patatin for result effective substitutions

Sequence	SEQ ID NO	Binding of modified peptides by patient IgE as measured by OD			
		HS03	HS06	HS01	HS02
DYFDVIGGTS	105		0.12	0.16	0.36
DYFDVIAGTS	106		0.14	0.17	0.4
VIGGTSSTGGL	107				0.04
VIAGTSTGAL	108				
AFYFEHGPQI	109		0.96	0.5	0.78
PAYFEHGPQI	110		0.75	0.41	0.69
PFAFEHGPQI	111				
PFYAEHGPQI	112		0.7	0.43	0.79
PFYFAHGPQI	113	0.93	1.07	0.59	1.44
PFYFEAGPQI	114	0.08	0.93	0.65	1.34
PFYFEHAPQI	115		0.75	0.54	1.11
PFYFEHGAQI	116		0.63	0.29	0.6
PFYFEHGPAT	117		0.63	0.25	0.56
PFYFEHGPQA	118		0.27	0.16	0.33
TFYLENGPKI	119	0.05	0.48	0.68	1.07
PFFFEGPQI	120				
AYLMQVLQEK	121		0.26	0.11	0.53
KALMQVLQEK	122				

KYAMQVLQEK	123	0.11	0.43	0.1	1.25
KYLAQVLQEK	124	0.22	0.48	0.11	1.34
KYLMAVLQEK	125	0.22	0.83	0.16	1.33
KYLMQALQEK	126	0.11	0.6	0.15	0.95
KYLMQVAQEK	127		0.53	0.15	0.81
KYLMQVLAEK	128	0.06	0.69	0.11	1.34
KYLMQVLQAK	129	0.74	0.79	0.05	0.58
KYLMQVLQEA	130		0.28	0.27	0.37
VFLHDKIKSL	131	0.06	0.26		0.41
AYSTAAAPTY	132		0.1	0.12	0.12
SASTAAAPTY	133				
SYATAAAPTY	134		0.16	0.13	0.37
SYSAAAAPTY	135		0.13	0.12	0.32
SYSTAAAATY	136		0.15	0.13	0.34
SYSTAAAPAY	137		0.15	0.14	0.29
SYSTAAAPTA	138		0.55	0.54	1.13
CISTSAAPTY	139	0.4			
SYSTAAAPAF	140	0.39	1.02	0.65	1.42
AFAAAAAPTY	141				0.07
SYSTAAAPTF	142	0.15	0.97	0.48	1.09
STSAAPTYFP	143		0.21	0.23	0.39
STSAAPTFFP	144				0.23
STSAAPTAFP	145				0.08
STAAAPTFFP	146			0.12	0.28
AAAATYFPPH	147		0.13	0.1	0.05
AAAPAYFPPH	148			0.07	0.04
AAAPTAFPPH	149				
AAAPTYAPPH	150		0.23	0.14	0.21
AAAPTYFAPH	151		0.45	0.18	0.44
AAAPTYFPAH	152		0.15	0.07	0.18
AAAPTYFPPA	153		0.1	0.06	0.31
SAAPTYFPAH	154		0.77	0.73	0.96
AAAPAFFPPH	155				
AAAPPFFPPH	156				
AAAPTFPPPH	157				
SISVATRLAQ	158			0.26	0.26
AMSLLTKEVH	159				
PAFASIRSLN	160				
PNFNAGSPTE	161				
KMLLLSLGTG	162				
NYLIISVGTG	163	0.49	1.08	0.64	1.48
KMLLLSLGAG	164		0.13		
AEFDKTYTAK	165	0.09	0.22		1.34
SAFDKTYTAK	166	0.66	0.71	0.06	1.42
SEADKTYTAK	167				0.99
SEFAKTYTAK	168	0.5	0.57		0.91
SEFDATYTAK	169				0.17
SEFDKAYTAK	170	0.1	0.24		1.38
SEFDKTATAK	171				0.81
SEFDKTYAAK	172	0.2	0.35		1.39
SEFDKTYTAA	173			0.1	1.18
KQAEKYTAEQ	174			0.08	0.24

SEFDAAFAAA	175				
SEFDKFTAK	176	0.09	0.16	0.07	1.45
AEKYTAEQCA	177				
ATYTAKEAAT	178		0.24		0.18
KAYTAKEAAT	179		0.28		0.33
KTATAKEAT	180				
KTYAAKEAAT	181	0.1	0.32		0.73
KTYAAAEAT	182				0.35
KTYTAKAAAT	183	0.4	0.59		0.82
KTYTAKEAAA	184				0.36
EKYTAEQCAK	185				
AAFAAAEAT	186				
KITTAKEAAT	187				
QALHCEKKYL	188				
QALDSKAAYL	189				
QALDSKNNFL	190				
QALHCENNLF	191				
CEKKYLRQD	192	1.01	0.16		
SKNNFLRVQE	193				
SENNYLRVQE	194	0.31	0.96	0.42	1
ALRVQENALT	195				
YARVQENALT	196	1.06	1.02	0.05	0.54
YLAVQENALT	197	0.37	1.04	0.11	1.06
YLRAQENALT	198	1.1	1	0.06	1.26
YL RVAENALT	199	1.03	0.92	0.08	1.26
YL RVQANALT	200	1.05	0.92	0.06	1.24
YL RVQEAALT	201	0.93	0.92	0.07	1.11
YL RVQENAAT	202	0.94	0.93	0.04	1.24
YL RVQENALA	203	1.05	0.96	0.43	1.16
YL RIQDDILT	204	1.07	0.85	0.39	1.12
YL TVAAAALT	205	1.05	0.86	0.28	1.33
FLRVQENALT	206				
NNYLRVQENA	207	0.23	0.88	0.5	1.17
KKYLRIQDDT	208		0.26	0.09	0.37
NNFLRVQENA	209				
NAYLRVQENA	210	0.17	1.02	0.53	1.06
ATYEEAKLRF	211	0.26	1.03		0.65
EAYEEAKLRF	212	0.06	0.43		0.33
ETAAEAKLRF	213		1.04		
ETYAEALKRF	214	0.62	1.02		1.15
ETYEAALKRF	215	1.06	0.38		0.89
ETYEEAAKRF	216	0.08	0.1		0.9
ETYEEALARF	217				0.11
ETYEEALKAF	218				0.1
ETYEEALKRA	219				0.1
GTNAQSLADF	220				
ETYEAALAAF	221	0.07	0.78	0.33	0.77
ETFEEALKRF	222				
YEEALKTFAK	223	1.08	0.85	0.14	1.46
FEEALKRFAK	224	0.46	0.72		0.67
AALKRFAKLL	225	0.15	0.17		
EAAKRFAKLL	226	0.08	0.33		0.05

EALARFAKLL	227	0.09		
EALKAFAKLL	228			
EALKRAAKLL	229	0.08	0.07	
EALKRFAALL	230			
EALKRFAKAL	231	0.06	0.09	0.1
EALKRFAKLA	232	0.06		0.1
QSLADFAKQL	233			
AALAAFAKLL	234			
LADFAKQLSD	235			
DFAKQLSDER	236			0.17
AFAALLSDRK	237			

1

2       Result effective substitutions were identified by a reduction in IgE binding ability  
 3 with respect to the non-substituted peptide sequence. Table 5 shows the identified result  
 4 effective substitutions. Blank spaces in the table indicate that binding of the peptide to  
 5 patient IgE was not detectable. Many substitutions of alanine or phenylalanine for the  
 6 original tyrosine resulted in reduced or eliminated antibody binding.

7

Table 5: Result effective substitutions of patatin

Location (SEQ ID NO)	Peptide (SEQ ID NO)	HS03	HS06	HS01	HS02
Minor Epitope 21	PFYFEHGPQI (36) ::A:::::::::: (111) ::F:::::::::: (r) (120) :::::::::::A (118)	1.08	0.17	0.65	0.06
Minor Epitope 27	KVLMQVLFQEK (42) :A:::::::::: (122) :::::::::::A (130) VFLHDKIKSL (c) (131)	1.08	0.79	0.12	0.04
Major Epitope 41	SYSTAAAPTY (56) A:::::::::: (132) :A:::::::::: (133) APFAA:::::::::: (r) (141) CI::S:::::::::: (c) (139)	1.04	0.55 0.1	1.15 0.12	0.25 0.12
Overlap Epitope 41/42	STAAAPTYFP (238) ::S:::::A:::: (r) (145)				0.08
Major Epitope 42 (57)	AAAPTYFPPH (57) :::::A:::::::::: (148) :::::A:::::::::: (149) :::::AF:::::::::: (r) (155) :::::PF:::::::::: (r) (156) :::::F:::::::::: (r) (157)	1.12	0.86	1.02 0.07	0.52 0.04
Major Epitope 61	SEFDKTYTAK (76) :::::A:::::::::: (169) KQAE:YTABQ (c) (174) :::::AAFA:A (r) (175)	0.12	0.28	1.33 0.08	1.35 0.17 0.24
Major	KTYTAKEAAT (77)	1.11	0.04	1.36	0.94

Epitope 62	A:::::::::: (178) ::A:::::::::: (180) :::::A:::::::::: (182) AAFA:A:::::::::: (r) (186) ::F:::::::::: (r) (187) EK::::EQC:::::: (c) (185)		0.24		0.18 0.35
Minor Epitope 72	QALDSKNNYL (87) :::HCEKK::: (c) (188) :::::AA::: (r) (189) ::::::::::F: (r) (190) :::::E::F: (r) (240)	0.28	0.43	0.44	0.46
Minor epitope 73	SKNNYLLRVQE (88) :::::F:::::::::: (r) (193)	1.31	0.21	0.44	0.05
Minor epitope 74	YLRVQENALT (87) A:::::::::: (195) F:::::::::: (r) (206)	1.31	0.2	1.38	0.03
Overlap epitope 73/74	NNYLRVQENA (207) ::F:::::::::: (r) (209)	0.23	0.88	0.5	1.17
Major epitope 85	ETYEEALKRF (100) ::::::::::A::: (217) ::::::::::A: (218) ::::::::::A (219) ::F:::::::::: (r) (222) G:NAQS:AD: (c) (220)	1.36	0.71	1.26	1.2 0.11 0.1 0.1
Major Epitope 86	EALKRFAKLL (101) :::::A:::::::::: (227) :::::A:::::::::: (228) ::::::::::A::: (229) ::::::::::A::: (230) ::::::::::A: (231) ::::::::::A (232) SD:AD::::Q: (c) (241) A::AA:::::::::: (r) (234)	0 0.08 0.06 0.06	1.12 0.09 0.07 0.09	1.38	0.04
Epitope overlap 86/87	LKRFAKLLSD (239) (NO BINDING)				
Major Epitope 87	RFAKLLSDRK (102) D::::Q::::ER (c) (236) A::A:::::::::: (r) (237)	0	0.94	0.98	0.05 0.17

1 (r) = rational; (c) = corn.

2 Example 5: Site directed mutagenesis

3 To introduce site specific mutations, the cloned DNA sequence of patatin (SEQ  
4 ID NO:1 encoding patatin protein SEQ ID NO:2; pMON 26820) was subjected to PCR  
5 with primers SEQ ID NO:3 and SEQ ID NO:4 to incorporate part of the  $\alpha$ -factor signal  
6 sequence (*Pichia* expression manual, Invitrogen, Carlsbad, CA), and EcoRI and Xhol  
7 restriction sites to facilitate cloning into the *Pichia pastoris* yeast secretion vector pPIC9  
8 (GenBank accession number Z46233; Invitrogen, Carlsbad, CA). Typical PCR

1 conditions are 25 cycles 94°C denaturation for 1 minute, 45°C annealing for one minute  
2 and 72°C extension for 2 minutes; plus one cycle 72°C extension for 10 minutes. A 50  
3  $\mu$ L reaction contains 30 pmol of each primer and 1  $\mu$ g of template DNA; and 1 X PCR  
4 buffer with MgCl<sub>2</sub>, 200  $\mu$ M dGTP, 200  $\mu$ M dATP, 200  $\mu$ M dTTP, 200  $\mu$ M dCTP, 2.5  
5 units of *Pwo* DNA polymerase. PCR reactions are performed in RoboCycler Gradient 96  
6 Temperature Cycler (Stratagene, La Jolla, CA).

7 The amplified fragment SEQ ID NO:5 was digested with restriction enzymes  
8 Xhol and EcoRI and cloned into the pBluescript vector (Stratagene, La Jolla, CA),  
9 digested with the same two restriction enzymes. The resulting plasmid (pMON 26869)  
10 was used for oligonucleotide-directed mutagenesis using the Bio-Rad mutagenesis kit  
11 based on the method of Kunkel (*Proc. Natl. Acad. Sci. U.S.A.*, 82: 477-492, 1985).  
12 Briefly, single-stranded pMON26869 was used as template for mutagenesis and was  
13 prepared by superinfection of plasmid containing cells with M13K07 (Gorman, *et al.*,  
14 *DNA Prot. Eng. Techniques*, 2: 3-10, 1990). The mutagenic oligonucleotides are SEQ ID  
15 NOS:8-15 (reverse complement). DNA purified from transformed DH5 $\alpha$  *E. coli* colonies  
16 was used for sequence determination. Sequencing was performed using the ABI PRISM  
17 sequencing kit (Perkin Elmer Biosystems, Foster City, CA). The resulting plasmid  
18 containing the mutation in the patatin gene was digested with restriction enzymes Xhol  
19 and EcoRI.

20 The patatin nucleic acid fragment was then ligated into the pPIC9 vector  
21 (Invitrogen, Carlsbad, CA), digested with the same two restriction enzymes to afford  
22 plasmid pMON37401. *Pichia pastoris* KM71 cells were electroporated with  
23 pMON37401 containing the appropriate mutation. The resulting transformed cells were  
24 used to produce protein in *Pichia pastoris* using the procedure supplied by the  
25 manufacturer (Invitrogen, Carlsbad, CA). The encoded protein contains an alpha factor  
26 signal cleavage site. Plasmid pMON37401 encodes SEQ ID NO:6 which is cleaved to  
27 afford SEQ ID NO:7, having four amino acids added at the N-terminus of amino acids  
28 24-386 of SEQ ID NO:2. Position four of SEQ ID NO:7 therefore corresponds to  
29 position 23 of SEQ ID NO:2.

30 The concentration of patatin in the culture was determined using a patatin ELISA  
31 assay and the enzyme activity was measured using the method of Hofgen and Willmitzer

1 (*Plant Science*, 66: 221-230, 1990). The variants containing multiple mutations were  
2 further purified using Mono Q and hydrophobic interaction chromatography (HIC). Each  
3 culture was purified by first sizing on Amicon YM10 membranes (Millipore, Bedford,  
4 MA) to a >10 kDa fraction, followed by chromatography on the Mono Q HR 10/10  
5 column (Pharmacia, Piscataway, NJ). For chromatography on the Mono Q column, the  
6 samples were loaded on the column in 25 mM Tris pH 7.5 and eluted with a gradient of  
7 1.0 M KCl in 25 mM Tris pH 7.5. Fractions containing patatin protein were determined  
8 using SDS-PAGE. For chromatography on the HIC column, the appropriate fractions  
9 were pooled and dialyzed into 1 M ammonium sulfate in 25 mM Tris pH 7.5. The  
10 dialyzed sample was then loaded on 16/10 phenyl Sepharose column (Pharmacia,  
11 Piscataway, NJ) and eluted with a gradient of 25 mM Tris pH 7.5.

12 The protein concentration was determined using the Bradford method, using BSA  
13 as a standard. SDS-PAGE analysis showed that these proteins were essentially pure. The  
14 esterase activity of the newly formed variants are shown in Table 6. The activity was  
15 determined using p-nitrophenyl caprate substrate as described by Hofgen and Willmitzer  
16 (*Plant Science*, 66: 221-230, 1990).

17 Table 6: Esterase activity of patatin mutants

Variant	Activity (mOD.min <sup>-1</sup> μg <sup>-1</sup> )
Wild type	93.2
Y106F	51.1
Y129F	74.7
Y185F	85.6
Y193F	82.2
Y185F/Y193F	99.4
Y270F	163.4
Y316F	94.88
Y362F	130.7
Y106F/Y129F/Y185F/Y193F/Y270F/Y316F/Y362F	57.1
Y185F/Y193F/Y270F/Y316F/Y362F	161.5

18  
19 Patatin proteins having a phenylalanine substitution at each of the amino acid  
20 positions 106, 129, 185, 193, 270, 316 and 362 (numbers correspond to positions in SEQ

1 ID NO:2) of expressed SEQ ID NO:7 exhibit full enzyme activity. Proteins having  
2 multiple substitutions also displayed full enzyme activity.

3 In addition to nucleotide sequences encoding conservative amino acid changes  
4 within the fundamental polypeptide sequence, biologically functional equivalent  
5 nucleotide sequences include nucleotide sequences containing other base substitutions,  
6 additions, or deletions. These include nucleic acids containing the same inherent genetic  
7 information as that contained in the cDNA which encode peptides, polypeptides, or  
8 proteins conferring pathogen resistance the same as or similar to that of pathogen upon  
9 host cells and plants. Such nucleotide sequences can be referred to as "genetically  
10 equivalent modified forms" of the cDNA, and can be identified by the methods  
11 described herein.

12 Mutations made in the cDNA, plasmid DNA, genomic DNA, synthetic DNA, or  
13 other nucleic acid encoding the deallergenized gene preferably preserve the reading  
14 frame of the coding sequence. Furthermore, these mutations preferably do not create  
15 complementary regions that could hybridize to produce secondary mRNA structures,  
16 such as loops or hairpins, that would adversely affect mRNA translation.

17 Although mutation sites can be predetermined, it is not necessary that the nature  
18 of the mutations *per se* be predetermined. For example, in order to select for optimum  
19 characteristics of mutants at a given site, random mutagenesis can be conducted at the  
20 target codon.

21 Alternatively, mutations can be introduced at particular loci by synthesizing  
22 oligonucleotides containing a mutant sequence, flanked by restriction sites enabling  
23 ligation to fragments of the native cDNA sequence. Following ligation, the resulting  
24 reconstructed nucleotide sequence encodes a derivative form having the desired amino  
25 acid insertion, substitution, or deletion.

26 **Example 6: Construction of permutein sequences**

27 Nucleic acid sequences encoding permutein proteins having rearranged N-  
28 terminus/C-terminus protein sequences can be made by following the general method  
29 described by Mullins et al. (*J. Am. Chem. Soc.* 116: 5529-5533, 1994). The steps are  
30 shown in Figure 3. The Figure and the following Examples involve the design and use of

1 a linker region separating the original C-terminus and N-terminus, but the use of a linker  
2 is not a critical or required element of permutein design.

3 Two sets of oligonucleotide primers are used in the construction of a nucleic acid  
4 sequence encoding a permutein protein. In the first step, oligonucleotide primers "new  
5 N-termini" and "linker start" are used in a PCR reaction to create amplified nucleic acid  
6 molecule "new N-termini fragment" that contains the nucleic acid sequence encoding the  
7 new N-terminal portion of the permutein protein, followed by the polypeptide linker that  
8 connects the C-terminal and N-terminal ends of the original protein. In the second step,  
9 oligonucleotide primers "new C-termini" and "linker end" are used in a PCR reaction to  
10 create amplified nucleic acid molecule "new C-termini fragment" that contains the  
11 nucleic acid sequence encoding the same linker as used above, followed by the new C-  
12 termini portion of the permutein protein. The "new N-termini" and "new C-termini"  
13 oligonucleotide primers are designed to include appropriate restriction enzyme  
14 recognition sites which assist in the cloning of the nucleic acid sequence encoding the  
15 permutein protein into plasmids.

16 Any suitable PCR conditions and polymerase can be used. It is desirable to use a  
17 thermostable DNA polymerase with high fidelity to reduce or eliminate the introduction  
18 of sequence errors. Typical PCR conditions are 25 cycles 94°C denaturation for 1  
19 minute, 45°C annealing for one minute and 72°C extension for 2 minutes; plus one cycle  
20 72°C extension for 10 minutes. A 50  $\mu$ L reaction contains 30 pmol of each primer and 1  
21  $\mu$ g of template DNA; and 1 X PCR buffer with MgCl<sub>2</sub>, 200  $\mu$ M dGTP, 200  $\mu$ M dATP,  
22 200  $\mu$ M dTTP, 200  $\mu$ M dCTP, 2.5 units of *Pwo* DNA polymerase. PCR reactions are  
23 performed in RoboCycler Gradient 96 Temperature Cycler (Stratagene, La Jolla, CA).

24 The amplified "new N-termini fragment" and "new C-termini fragment" are  
25 annealed to form a template in a third PCR reaction to amplify the full-length nucleic  
26 acid sequence encoding the permutein protein. The DNA fragments "new N-termini  
27 fragment" and "new C-termini fragment" are resolved on a 1% TAE gel, stained with  
28 ethidium bromide, and isolated using the QIAquick Gel Extraction Kit (Qiagen, Valencia,  
29 CA). These fragments are combined in equimolar quantities with oligonucleotide  
30 primers "new N-termini" and "new C-termini" in the third PCR reaction. The conditions

1 for the PCR are the same as used previously. PCR reaction products can be purified  
2 using the QIAquick PCR purification kit (Qiagen, Valencia, CA).

3 Alternatively, a linker sequence can be designed containing a restriction site,  
4 allowing direct ligation of the two amplified PCR products.

5 Example 7: Construction of plasmid pMON 37402

6 The patatin protein contains a trypsin protease sensitive site at the arginine amino  
7 acid at position 246, as determined by electrophoresis of a trypsin digest reaction. In  
8 order to determine if the exposed protease site is an antigenic epitope, a permutein was  
9 constructed using positions 246-247 as a breakpoint.

10 The nucleic acid sequence encoding the permutein protein in plasmid pMON  
11 37402 was created using the method illustrated in Figure 3 and described in Example 6.  
12 Nucleic acid molecule “new N-termini fragment” was created and amplified from the  
13 sequence encoding patatin in plasmid pMON26820 using oligonucleotide primers 27  
14 (SEQ ID NO:242) and 48 (SEQ ID NO:243). Nucleic acid molecule “new C-termini  
15 fragment” was created and amplified from the sequence encoding patatin in plasmid  
16 pMON26820 using oligonucleotide primers 47 (SEQ ID NO:244) and 36 (SEQ ID  
17 NO:245). The full-length nucleic acid molecule encoding the permutein protein was  
18 created and amplified from annealed fragments “new N-termini fragment” and “new C-  
19 termini fragment” using oligonucleotide primers 27 (SEQ ID NO:242) and 36 (SEQ ID  
20 NO:245).

21 The resulting amplified nucleic acid molecule was digested with restriction  
22 endonucleases Xhol and EcoRI, and purified using the QIAquick PCR purification kit  
23 (Qiagen, Valencia, CA). Plasmid pMON 26869 (derivative of pPIC9, Invitrogen,  
24 Carlsbad, CA) was digested with restriction endonucleases Xhol and EcoRI, and gel  
25 purified, resulting in an approximately 2900 base pair vector fragment. The purified  
26 restriction fragments were combined and ligated using T4 DNA ligase.

27 The ligation reaction mixture was used to transform *E. coli* strain DH5 $\alpha$  cells  
28 (Life Technologies, Gaithersburg, MD). Transformant bacteria were selected on  
29 ampicillin-containing plates. Plasmid DNA was isolated and sequenced to confirm the

1 presence of the correct insert. The resulting plasmid was designated pMON 37402  
2 (containing SEQ ID NO:246, encoding protein sequence SEQ ID NO:247).

3 Example 8: Construction of plasmid pMON 37405

4 Amino acids 201-202, near tyrosine 193, were chosen as a breakpoint for the  
5 construction of a permutein protein.

6 The nucleic acid sequence encoding the permutein protein in plasmid pMON  
7 37405 was created using the method illustrated in Figure 3 and described in Example 6.  
8 Nucleic acid molecule “New N-termini fragment” was created and amplified from the  
9 sequence encoding patatin in plasmid pMON26820 using oligonucleotide primers 48  
10 (SEQ ID NO:243) and 58 (SEQ ID NO:249). Nucleic acid molecule “New C-termini  
11 fragment” was created and amplified from the sequence encoding patatin in plasmid  
12 pMON26820 using oligonucleotide primers 47 (SEQ ID NO:244) and 59 (SEQ ID  
13 NO:249). The full-length nucleic acid molecule encoding the permutein protein was  
14 created and amplified from annealed fragments “New N-termini fragment” and “New C-  
15 termini fragment” using oligonucleotide primers 58 (SEQ ID NO:248) and 59 (SEQ ID  
16 NO:249).

17 The resulting amplified nucleic acid molecule was digested with restriction  
18 endonucleases XhoI and EcoRI, and purified using the QIAquick PCR purification kit  
19 (Qiagen, Valencia, CA). Plasmid pMON 26869 (derivative of pPIC9, Invitrogen,  
20 Carlsbad, CA) was digested with restriction endonucleases XhoI and EcoRI, and gel  
21 purified, resulting in an approximately 2900 base pair vector fragment. The purified  
22 restriction fragments were combined and ligated using T4 DNA ligase.

23 The ligation reaction mixture was used to transform *E. coli* strain DH5 $\alpha$  cells  
24 (Life Technologies, Gaithersburg, MD). Transformant bacteria were selected on  
25 ampicillin-containing plates. Plasmid DNA was isolated and sequenced to confirm the  
26 presence of the correct insert. The resulting plasmid was designated pMON 37405  
27 (containing SEQ ID NO:250, encoding protein sequence SEQ ID NO:251).

1                   Example 9: Construction of plasmid pMON 37406

2                   Amino acids 183-184, adjacent to tyrosine 185, were chosen as a breakpoint for  
3                   the construction of a permutein protein.

4                   The nucleic acid sequence encoding the permutein protein in plasmid pMON  
5                   37406 was created using the method illustrated in Figure 3 and described in Example 6.  
6                   Nucleic acid molecule “New N-termini fragment” was created and amplified from the  
7                   sequence encoding patatin in plasmid pMON26820 using oligonucleotide primers 48  
8                   (SEQ ID NO:243) and 60 (SEQ ID NO:252). Nucleic acid molecule “New C-termini  
9                   fragment” was created and amplified from the sequence encoding patatin in plasmid  
10                   pMON26820 using oligonucleotide primers 47 (SEQ ID NO:244) and 61 (SEQ ID  
11                   NO:253). The full-length nucleic acid molecule encoding the permutein protein was  
12                   created and amplified from annealed fragments “New N-termini fragment” and “New C-  
13                   termini fragment” using oligonucleotide primers 60 (SEQ ID NO:252) and 61 (SEQ ID  
14                   NO:253).

15                   The resulting amplified nucleic acid molecule was digested with restriction  
16                   endonucleases XhoI and EcoRI, and purified using the QIAquick PCR purification kit  
17                   (Qiagen, Valencia, CA). Plasmid pMON 26869 (derivative of pPIC9, Invitrogen,  
18                   Carlsbad, CA) was digested with restriction endonucleases XhoI and EcoRI, and gel  
19                   purified, resulting in an approximately 2900 base pair vector fragment. The purified  
20                   restriction fragments were combined and ligated using T4 DNA ligase.

21                   The ligation reaction mixture was used to transform *E. coli* strain DH5 $\alpha$  cells  
22                   (Life Technologies, Gaithersburg, MD). Transformant bacteria were selected on  
23                   ampicillin-containing plates. Plasmid DNA was isolated and sequenced to confirm the  
24                   presence of the correct insert. The resulting plasmid was designated pMON37406  
25                   (containing SEQ ID NO:254, encoding protein sequence SEQ ID NO:255).

26                   Example 10: Construction of plasmid pMON 37407

27                   Amino acids 268-269, adjacent to tyrosine 270, were chosen as a breakpoint for  
28                   the construction of a permutein protein.

1        The nucleic acid sequence encoding the permutein protein in plasmid pMON  
2 37407 was created using the method illustrated in Figure 3 and described in Example 6.  
3 Nucleic acid molecule “New N-termini fragment” was created and amplified from the  
4 sequence encoding patatin in plasmid pMON26820 using oligonucleotide primers 48  
5 (SEQ ID NO:243) and 62 (SEQ ID NO:256). Nucleic acid molecule “New C-termini  
6 fragment” was created and amplified from the sequence encoding patatin in plasmid  
7 pMON26820 using oligonucleotide primers 47 (SEQ ID NO:244) and 63 (SEQ ID  
8 NO:257). The full-length nucleic acid molecule encoding the permutein protein was  
9 created and amplified from annealed fragments “New N-termini fragment” and “New C-  
10 termini fragment” using oligonucleotide primers 62 (SEQ ID NO:256) and 63 (SEQ ID  
11 NO:257).

12       The resulting amplified nucleic acid molecule was digested with restriction  
13 endonucleases XhoI and EcoRI, and purified using the QIAquick PCR purification kit  
14 (Qiagen, Valencia, CA). Plasmid pMON 26869 (derivative of pPIC9, Invitrogen,  
15 Carlsbad, CA) was digested with restriction endonucleases XhoI and EcoRI, and gel  
16 purified, resulting in an approximately 2900 base pair vector fragment. The purified  
17 restriction fragments were combined and ligated using T4 DNA ligase.

18       The ligation reaction mixture was used to transform *E. coli* strain DH5 $\alpha$  cells  
19 (Life Technologies, Gaithersburg, MD). Transformant bacteria were selected on  
20 ampicillin-containing plates. Plasmid DNA was isolated and sequenced to confirm the  
21 presence of the correct insert. The resulting plasmid was designated pMON37407  
22 (containing SEQ ID NO:258, encoding protein sequence SEQ ID NO:259).

23       Example 11: Construction of plasmid pMON 37408

24       Amino acids 321-322, near tyrosine 216, were chosen as a breakpoint for the  
25 construction of a permutein protein.

26       The nucleic acid sequence encoding the permutein protein in plasmid pMON  
27 37408 was created using the method illustrated in Figure 3 and described in Example 6.  
28 Nucleic acid molecule “New N-termini fragment” was created and amplified from the  
29 sequence encoding patatin in plasmid pMON26820 using oligonucleotide primers 48  
30 (SEQ ID NO:243) and 64 (SEQ ID NO:260). Nucleic acid molecule “New C-termini

1 fragment" was created and amplified from the sequence encoding patatin in plasmid  
2 pMON26820 using oligonucleotide primers 47 (SEQ ID NO:244) and 65 (SEQ ID  
3 NO:261). The full-length nucleic acid molecule encoding the permutein protein was  
4 created and amplified from annealed fragments "New N-termini fragment" and "New C-  
5 termini fragment" using oligonucleotide primers 64 (SEQ ID NO:260) and 65 (SEQ ID  
6 NO:261).

7 The resulting amplified nucleic acid molecule was digested with restriction  
8 endonucleases Xhol and EcoRI, and purified using the QIAquick PCR purification kit  
9 (Qiagen, Valencia, CA). Plasmid pMON 26869 (derivative of pPIC9, Invitrogen,  
10 Carlsbad, CA) was digested with restriction endonucleases XhoI and EcoRI, and gel  
11 purified, resulting in an approximately 2900 base pair vector fragment. The purified  
12 restriction fragments were combined and ligated using T4 DNA ligase.

13 The ligation reaction mixture was used to transform *E. coli* strain DH5 $\alpha$  cells  
14 (Life Technologies, Gaithersburg, MD). Transformant bacteria were selected on  
15 ampicillin-containing plates. Plasmid DNA was isolated and sequenced to confirm the  
16 presence of the correct insert. The resulting plasmid was designated pMON37408  
17 (containing SEQ ID NO:262, encoding protein sequence SEQ ID NO:263).

18 Example 12: Production of permutein proteins in *Pichia pastoris*

19 Plasmids pMON37402, pMON37405, pMON37406, pMON37407, and  
20 pMON37408 were individually used to electroporate KM71 cells from *Pichia pastoris*  
21 according to the procedure supplied by the manufacturer (Invitrogen, Carlsbad, CA). The  
22 resulting transformed cells were used to produce protein in *Pichia pastoris* following the  
23 procedure supplied by the manufacturer (Invitrogen, Carlsbad, CA).

24 The concentration of patatin in the culture was determined using a patatin ELISA  
25 assay and the enzyme activity was measured using the method of Hofgen and Willmitzer  
26 (*Plant Science*, 66: 221-230, 1990). The variants containing multiple mutations were  
27 further purified using Mono Q and hydrophobic interaction chromatography (HIC). Each  
28 culture was purified by first sizing on YM10 membranes (Amicon, MA) to a [>10 kDa]  
29 fraction, followed by chromatography on the Mono Q HR 10/10 column (Pharmacia, NJ).  
30 For chromatography on the Mono Q column, the samples were loaded on the column in

1 25 mM Tris pH 7.5 and eluted with a gradient of 1.0 M KCl in 25 mM Tris pH 7.5.  
2 Fractions containing patatin protein were determined using SDS-PAGE. For  
3 chromatography on the HIC column, the appropriate fractions were pooled and dialyzed  
4 into 1 M ammonium sulfate in 25 mM Tris pH 7.5. The dialyzed sample was then loaded  
5 on 16/10 phenyl Sepharose column (Pharmacia, NJ) and eluted with a gradient of 25 mM  
6 Tris pH 7.5.

7 The protein concentration was determined using the Bradford method, using BSA  
8 as a standard. SDS-PAGE analysis showed that these proteins were essentially pure. The  
9 esterase activity of the variants are shown in Table 7.

10 Table 7: Activity of permuteins

pMON	Breakpoint	Activity ( $\Delta\text{OD min}^{-1} \mu\text{g}^{-1}$ )
Native enzyme		83.21
pMON37402	246/247	66.7
pMON37405	201/202	No expression
pMON37406	183/184	No expression
pMON37407	268/269	12.1
pMON37408	321/322	No expression

11  
12 The activity was determined using *p*-nitrophenyl caprate substrate as described by  
13 Hofgen and Willmitzer (*Plant Science*, 66: 221-230, 1990).

14 Example 13: Insect bioefficacy assays

15 Assays for activity against larvae of SCRW are carried out by overlaying the test  
16 sample on an agar diet similar to that described by Marrone (*J. Econ. Entom.* 78: 290-  
17 293, 1985). Test samples were prepared in 25 mM Tris, pH 7.5 buffer. Neonate larvae  
18 are allowed to feed on the treated diet at 26°C, and mortality and growth stunting were  
19 evaluated after 5 or 6 days. The results of this assay are shown in Table 8.

20 Table 8: Insect bioefficacy assay

Protein (200 ppm)	Mean Survival Weight	% Weight Reduction
Tris buffer (control)	1.26 ± 0.3	-
Wild Type	0.21 ± 0.02	83

pMON37402	0.21 ± 0.03	83
pMON37407	0.32 ± 0.04	75

1

2 These data demonstrate that the growth of the SCRW larvae is similarly reduced  
 3 upon ingestion of the proteins encoded by pMON37402 and pMON37407 as compared to  
 4 the wild type patatin protein.

5 Example 14: Permutein sequences improved for monocot expression

6 Modification of coding sequences has been demonstrated above to improve  
 7 expression of insecticidal proteins. A modified coding sequence was thus designed to  
 8 improve expression in plants, especially corn (SEQ ID NO:264).

9 Example 15: Construction of pMON40701 for monocot expression

10 Plasmid pMON19767 was digested with restriction endonucleases NcoI and  
 11 EcoRI and the 1100 bp gene fragment was purified using the QIAquick PCR purification  
 12 kit (Qiagen, Valencia, CA). Plasmid pMON33719 was digested with restriction  
 13 endonucleases NcoI and EcoRI, and gel purified, resulting in an approximately 3900 base  
 14 pair vector fragment. The two purified restriction fragments were combined and ligated  
 15 using T4 DNA ligase.

16 The ligation reaction mixture was used to transform *E. coli* strain DH5 $\alpha$  cells  
 17 (Life Technologies, Gaithersburg, MD). Transformant bacteria were selected on  
 18 ampicillin-containing plates. Plasmid DNA was isolated and sequenced to confirm the  
 19 presence of the correct insert. The resulting plasmid was designated pMON40700.  
 20 Plasmid pMON40700 was digested with restriction endonuclease NotI and the resulting  
 21 2200 bp DNA fragment was purified using the QIAquick PCR purification kit (Qiagen,  
 22 Valencia, CA). Plasmid pMON30460 was digested with restriction endonuclease NotI,  
 23 and gel purified, resulting in an approximately 4200 base pair vector fragment. The two  
 24 purified restriction fragments were combined and ligated using T4 DNA ligase.

25 The ligation reaction mixture was used to transform *E. coli* strain DH5 $\alpha$  cells  
 26 (Life Technologies, Gaithersburg, MD). Transformant bacteria were selected on

1 kanamycin-containing plates. The resulting plasmid was designated pMON40701  
2 (containing SEQ ID NO:264, encoding protein sequence SEQ ID NO:265).

3 Example 16: Construction of pMON40703 for monocot expression

4 The nucleic acid sequence encoding the permutein protein in plasmid  
5 pMON40703 was created using the method illustrated in Figure 3 and described in  
6 Example 6. Nucleic acid molecule “New N-termini fragment” was created and amplified  
7 from the sequence encoding patatin in plasmid pMON19767 using oligonucleotide  
8 primers Syn1 (SEQ ID NO:266) and Syn2 (SEQ ID NO:267). Nucleic acid molecule  
9 “New C-termini fragment” was created and amplified from the sequence encoding patatin  
10 in plasmid pMON19767 using oligonucleotide primers Syn3 (SEQ ID NO:268) and Syn4  
11 (SEQ ID NO:269). The full-length nucleic acid molecule encoding the permutein protein  
12 was created and amplified from annealed fragments “New N-termini fragment” and  
13 “New C-termini fragment” using oligonucleotide primers Syn1 (SEQ ID NO:266) and  
14 Syn4 (SEQ ID NO:269).

15 The resulting amplified nucleic acid molecule was digested with restriction  
16 endonucleases NcoI and EcoRI, and purified using the QIAquick PCR purification kit  
17 (Qiagen, Valencia, CA). Plasmid pMON33719 was digested with restriction  
18 endonucleases NcoI and EcoRI, and gel purified, resulting in an approximately 3900 base  
19 pair vector fragment. The purified restriction fragments were combined and ligated using  
20 T4 DNA ligase.

21 The ligation reaction mixture was used to transform *E. coli* strain DH5 $\alpha$  cells  
22 (Life Technologies, Gaithersburg, MD). Transformant bacteria were selected on  
23 ampicillin-containing plates. Plasmid DNA was isolated and sequenced to confirm the  
24 presence of the correct insert. The resulting plasmid was designated pMON40702.  
25 Plasmid pMON40702 was digested with NotI, and the resulting 2200 bp DNA fragment  
26 was purified using the QIAquick PCR purification kit (Qiagen, Valencia, CA). Plasmid  
27 pMON30460 was digested with restriction endonuclease NotI, and gel purified, resulting  
28 in an approximately 4200 base pair vector fragment. The purified restriction fragments  
29 were combined and ligated using T4 DNA ligase.

1        The ligation reaction mixture was used to transform *E. coli* strain DH5 $\alpha$  cells  
2 (Life Technologies, Gaithersburg, MD). Transformant bacteria were selected on  
3 kanamycin-containing plates. The resulting plasmid was designated pMON40703  
4 (containing SEQ ID NO:270, encoding protein sequence SEQ ID NO:271). Plasmid  
5 pMON40703 encodes a permutein protein with a “breakpoint” at positions 246/247 of the  
6 wild type patatin protein sequence (SEQ ID NO:2). The first 23 amino acids of SEQ ID  
7 NO:2 are a signal peptide sequence which is cleaved in the mature protein.

8        Example 17: Construction of pMON40705 for monocot expression

9        The nucleic acid sequence encoding the permutein protein in plasmid  
10 pMON40705 was created using the method illustrated in Figure 3 and described in  
11 Example 6. Nucleic acid molecule “New N-termini fragment” was created and amplified  
12 from the sequence encoding patatin in plasmid pMON19767 using oligonucleotide  
13 primers Syn10 (SEQ ID NO:272) and Syn2 (SEQ ID NO:267). Nucleic acid molecule  
14 “New C-termini fragment” was created and amplified from the sequence encoding patatin  
15 in plasmid pMON19767 using oligonucleotide primers Syn3 (SEQ ID NO:268) and  
16 Syn11 (SEQ ID NO:273). The full-length nucleic acid molecule encoding the permutein  
17 protein was created and amplified from annealed fragments “New N-termini fragment”  
18 and “New C-termini fragment” using oligonucleotide primers Syn10 (SEQ ID NO:272)  
19 and Syn11 (SEQ ID NO:273).

20        The resulting amplified nucleic acid molecule was digested with restriction  
21 endonucleases NcoI and EcoRI, and purified using the QIAquick PCR purification kit  
22 (Qiagen, Valencia, CA). Plasmid pMON33719 was digested with restriction  
23 endonucleases NcoI and EcoRI, and gel purified, resulting in an approximately 3900 base  
24 pair vector fragment. The purified restriction fragments were combined and ligated using  
25 T4 DNA ligase.

26        The ligation reaction mixture was used to transform *E. coli* strain DH5 $\alpha$  cells  
27 (Life Technologies, Gaithersburg, MD). Transformant bacteria were selected on  
28 ampicillin-containing plates. Plasmid DNA was isolated and sequenced to confirm the  
29 presence of the correct insert. The resulting plasmid was designated pMON40704.  
30 Plasmid pMON40704 was digested with restriction endonuclease NotI, and the resulting

1 2200 bp DNA fragment was purified using the QIAquick PCR purification kit (Qiagen,  
2 Valencia, CA). Plasmid pMON30460 was digested with restriction endonuclease NotI,  
3 and gel purified, resulting in an approximately 4200 base pair vector fragment. The  
4 purified restriction fragments were combined and ligated using T4 DNA ligase.

5 The ligation reaction mixture was used to transform *E. coli* strain DH5 $\alpha$  cells  
6 (Life Technologies, Gaithersburg, MD). Transformant bacteria were selected on plates  
7 containing kanamycin. The resulting plasmid was designated pMON40705 (containing  
8 SEQ ID NO:274, encoding protein sequence SEQ ID NO:275). Plasmid pMON40705  
9 encodes a permutein protein with a “breakpoint” at positions 268/269 of the wild type  
10 patatin protein sequence (SEQ ID NO:2). The first 23 amino acids of SEQ ID NO:2 are a  
11 signal peptide sequence which is cleaved in the mature protein.

12 Example 18: Transient expression of protein in corn leaf protoplasts

13 Plasmids pMON40701, pMON40703, and pMON40705 (all containing the native  
14 signal sequence for vacuolar targeting) were separately electroporated into corn leaf  
15 protoplasts as described by Sheen (*Plant Cell* 3: 225-245, 1991). Protein was extracted  
16 with glass beads and the supernatant was assayed for protein expression using ELISA for  
17 patatin and NPTII. Expression of protein by the transformed corn protoplasts was  
18 confirmed by Western blot analysis. Expression results are shown in Table 9.

19 Table 9: ELISA data

Sample	Patatin ELISA ( $\mu$ g/mL)	NPTII ELISA ( $\mu$ g/mL)	Normalized Expression (Patatin ELISA/NPTII ELISA)
pMON40701	1.1	0.6	1.8
pMON40703	2.1	0.3	7.0
pMON40705	1.3	0.6	2.2

20 The results indicate that the permutein encoded by plasmid pMON40703  
21 surprisingly shows approximately 4-fold higher expression compared to the wild type  
22 enzyme.

1      Example 19: Deglycosylation of protein sequences

2      This example provides evidence that glycosylation of can contribute to the  
3      allergenicity of a protein. Accordingly, rational substitution of amino acid residues likely  
4      to be the targets of glycosylation within a subject allergen protein may reduce or  
5      eliminate the allergenic properties of the protein without adversely affecting the  
6      enzymatic, insecticidal, antifungal or other functional properties of the protein.

7      Glycosylation commonly occurs as either N-linked or O-linked forms. N-linked  
8      glycosylation usually occurs at the motif Asn-Xaa-Ser/Thr, where Xaa is any amino acid  
9      except Pro (Kasturi, L. et al., *Biochem J.* 323: 415-519, 1997; Melquist, J.L. et al.,  
10     *Biochemistry* 37: 6833-6837, 1998). O-linked glycosylation occurs between the hydroxyl  
11     group of serine or threonine and an amino sugar.

12     Site directed mutagenesis of selected asparagine, serine, or threonine may be used  
13     to reduce or eliminate the glycosylation of patatin proteins. A search of SEQ ID NO:2  
14     for the Asn-Xaa-Ser/Thr motif reveals one occurrence at amino acid positions 202-204.  
15     Mutagenization of the nucleic acid sequence encoding this region results in a reduced  
16     allergenicity of the encoded protein.

17     In order to test this approach to reducing allergenicity of patatin proteins, two sets  
18     of experiments were performed: a) production of patatin proteins in *Escherichia coli*,  
19     which do not glycosylate proteins; and b) production of patatin proteins with an N202Q  
20     site directed mutation.

21     Antibodies obtained from patients HS-07 and G15-MON (not potato allergic) did  
22     not show specific binding to wild type patatin, patatin produced in *E. coli*, or the N202Q  
23     variant. Antibodies obtained from patient HS-01 (potato allergic) bound to wild type  
24     patatin, but not to patatin produced in *E. coli* or the N202Q variant. Antibodies obtained  
25     from patient HS-02 (potato allergic) bound strongly to wild type patatin, but extremely  
26     weakly to patatin produced in *E. coli*, and binding to the N202Q variant resembled vector  
27     controls. Antibodies obtained from patient HS-03 (potato allergic) bound to wild type  
28     patatin, but not to patatin produced in *E. coli* or the N202Q variant. Antibodies obtained  
29     from patient HS-05 (potato allergic) bound to wild type patatin, but very weakly to  
30     patatin produced in *E. coli* and the N202Q variant. Antibodies obtained from patient HS-  
31     06 (potato allergic) strongly bound wild type patatin, the N202Q variant, and to patatin

1 produced in *E. coli*. These results strongly suggest that glycosylation is at least partially  
2 responsible for the antigenic properties of patatin proteins, and that site directed  
3 mutagenesis may be used to reduce or eliminate specific antibody binding. Mutagenesis  
4 at position 202 of SEQ ID NO:2 may be useful for reducing or eliminating specific  
5 antibody binding.

6 The deglycosylation approach was also tested using a patatin homolog, Pat17. As  
7 demonstrated above, patatin epitopes exhibiting IgE binding were identified, and each  
8 contained a Tyr residue. Substitution of these Tyr residues within each epitope led to loss  
9 of IgE binding. Site-directed mutagenesis was used to produce variants with individual  
10 and multiple Tyr substitutions in the protein, which was expressed in *Pichia pastoris* and  
11 assessed for enzyme activity. All the variants were found to have enzymatic activity no  
12 less than the wild type protein. A single variant with all 5 tyrosine residues substituted  
13 with phenylalanine was found to have insecticidal activity no less than the unsubstituted  
14 protein and was expressed in *E.coli* to produce the non-glycosylated version. The *E.coli*  
15 5-"Tyr to Phe" variant was assessed for IgE binding. An isozyme of patatin, designated  
16 Pat17, was also expressed in corn to produce a plant glycoprotein and in *E.coli* to  
17 produce a nonglycosylated protein. Sera of seven patients (five exhibiting potato allergy  
18 and one exhibiting other allergies but no allergy to potatoes) were used to assay  
19 Pat17 or Pat17 variant binding by immunoblot assay. Four of the five sera from patients  
20 exhibiting potato allergy showed either very weak or no binding to wild type patatin  
21 expressed in *E.coli* but did bind to the 5-Tyr variant. Serum from one patient exhibiting  
22 potato allergy showed strong binding to recombinant wild type patatin protein expressed  
23 in *E.coli* but weak binding to the 5-Tyr variant. Sera from all five patients exhibiting  
24 potato allergy bound strongly to patatin expressed in corn and native patatin present in  
25 potatoes. Serum from a control patient allergic to eggs, milk, peanuts and seafood, but  
26 exhibiting no allergy to potatoes showed no binding to patatin expressed in *E.coli* but did  
27 bind to patatin expressed in corn. Immunoblot results suggested that the sugar moiety in  
28 patatin is a non-specific IgE binding epitope and the polypeptide portion of patatin also  
29 contains immunogenic IgE epitopes.

1 Patients who suffer from potato allergy were identified at Johns Hopkins Clinic  
2 (Baltimore, MD) and were evaluated for potato allergy using clinical criteria outlined in  
3 Table 2.

4 Serum was obtained from patients with convincing clinical history of potato  
5 allergy. The convincing history was defined as being one or more of the following: a)  
6 positive potato allergic reaction as evaluated by double-blind placebo-control food  
7 challenge b) anaphylaxis and/or hospitalization due to the consumption of potatoes or c)  
8 dramatic skin test results.

9 Peptide Synthesis

10 Peptides were synthesized on cellulose membranes using the SPOTS system  
11 (Genosys Biotechnologies, TX). Membranes were stored at -20°C until use.

12 Site directed Mutagenesis

13 Site specific mutations were introduced into patatin by first incorporating part of  
14 the a-factor signal sequence (*Pichia* expression manual, Invitrogen, Carlsbad, CA) to the  
15 patatin gene using PCR. Primers used for the PCR were  
16 GGAGCTCGAGAAAAGAGAGGGCTGAAGCTCAGTTGGGAGAAATGGTGAATGT  
17 TCT (*Xba*I site in italics) and GGTCTAGAG *GAATTCTCATTAATAAGAAG* (*Eco*RI  
18 site in italics). The primers contained restriction sites to facilitate cloning into *Pichia*  
19 *pastoris* yeast secretion vector pPIC9 (GenBank accession number Z46233; Invitrogen,  
20 Carlsbad, CA). Typical PCR conditions are 25 cycles 94°C denaturation for 1 minute,  
21 45°C annealing for one minute and 72°C extension for 2 minutes; plus one cycle 72°C  
22 extension for 10 minutes. A 50 mL reaction contained 30 pmol of each primer and 1 mg  
23 of template DNA; and 1 X PCR buffer with MgCl<sub>2</sub>, 200 mM dGTP, 200 mM dATP, 200  
24 mM dTTP, 200 mM dCTP, 2.5 units of *Pwo* DNA polymerase. PCR reactions are  
25 performed in RoboCycler Gradient 96 Temperature Cycler (Stratagene, La Jolla, CA).

26 The amplified patatin gene was digested with restriction enzymes *Xba*I and *Eco*RI  
27 and cloned into the pBluescript vector (Stratagene, La Jolla, CA), digested with the same  
28 two restriction enzymes. The template plasmid DNA used for the PCR was  
29 pMON26820. The resulting plasmid (pMON 26869) was used for oligonucleotide-  
30 directed mutagenesis using the Bio-Rad mutagenesis kit based on the method of Kunkel

1 et al., *Proc Natl Acad Sci USA* 82, 477-92 (1985). Briefly, single-stranded pMON26869  
2 was used as template for mutagenesis and was prepared by superinfection of plasmid  
3 containing cells with M13K07 (Gorman et al., *DNA and Protein Engineering techniques*  
4 2, 3-10 (1990)). DNA purified from transformed DH5a *E. coli* colonies was used for  
5 sequence determination. Sequencing was performed using the ABI PRISM sequencing  
6 kit (Perkin Elmer Biosystems, Foster City, CA).

7 Protein Expression in *Pichia pastoris*

8 Plasmids containing the mutations in the patatin gene were digested with restriction  
9 enzymes *Xba*I and *Eco*RI. The patatin nucleic acid fragment was then ligated into the  
10 pPIC9 vector (Invitrogen, Carlsbad, CA), digested with the same two restriction enzymes  
11 to afford plasmid pMON37401. *Pichia pastoris* KM71 cells were electroporated with  
12 pMON37401 containing the appropriate mutation. The resulting transformed cells were  
13 used to produce protein in *Pichia pastoris* using the procedure supplied by the  
14 manufacturer (Invitrogen, Carlsbad, CA). The proteins were purified in the same way as  
15 the proteins expressed in *E. coli* (see below).

16 Western Blotting of Proteins

17 Protein samples were electrophoresed by SDS-PAGE and electroblotted onto PVDF  
18 membrane (Millipore, Bedford MA). Protein blots were processed by standard Western  
19 blotting (immunoblotting) techniques and were incubated in potato allergic serum diluted  
20 1:5 in PBS buffer for 1 hour. After washing the blots 3 times with PBS, the blots were  
21 incubated in biotinylated anti-IgE (Johns Hopkins Hospital, Baltimore MD) for 1 hour,  
22 followed by a 30 minute incubation in HRP-linked avidin (Promega, New York, NY).  
23 IgE-reactive protein bands were visualized by using the ECL system (Amersham  
24 Pharmacia Biotech, NJ). As a control, one blot was incubated in anti-IgE only. His-  
25 tagged glyphosate oxidase and potato extracts was prepared and provided for this study  
26 by Regulatory Sciences, Monsanto Company. The peptides were evaluated using the  
27 same incubation procedures as described above.

28 Expression and purification of patatin in corn

29 An isozyme of patatin, Pat17, was generated for expression in corn using a modified  
30 plant optimized gene sequence as described by Brown *et al* (US Patent 5,689,052). All

1 the constructs contained the native 23 amino acid signal peptide for vacuolar targeting.  
2 Corn was transformed by microprojectile bombardment (Morrish et al., in *Transgenic*  
3 *plants. Fundamentals and Applications* (ed. Hiatt, A.) 133-171 (Marcel Dekker, New  
4 York, 1993); Songstad et al., *In Vitro Cell Dev Biol - Plant* 32, 179-183 (1996)).  
5 Protein from the transformed corn plants was purified by first grinding the leaves in  
6 liquid nitrogen and extracting the protein using 25 mM Tris/HCl. The plant extract was  
7 further dialyzed against 25 mM Tris/HCl pH 7.5. The plant extract was then loaded onto  
8 Mono Q HR 10/10 anion-exchange column (Amersham Pharmacia, NJ) equilibrated with  
9 25 mM Tris/HCl pH 7.5 (buffer A). The protein was eluted with 25 mM Tris/HCl pH  
10 7.5, 1 M KCl (buffer B) using a linear gradient of 0-100% buffer B using an HPLC  
11 system (Shimadzu). Fractions containing protein were assayed for esterase activity and  
12 dialyzed against 25 mM Tris/HCl pH 7.5, 1 M Ammonium Sulfate (buffer C). The  
13 protein was purified to homogeneity by loading onto a phenyl-Sepharose 16/10 column  
14 (Amersham Pharmacia, NJ) equilibrated with buffer C. Esterase active fractions were  
15 pooled and dialyzed against 25 mM Tris/HCl pH 7.5.

16 Expression and purification of patatin in *E.coli*

17 Pat17 was expressed in *E.coli* using the pET expression system (Novagen, WI). The  
18 coding region of the mature Pat17 gene (without its signal peptide) was amplified by  
19 PCR using the primers 5'-GGGCCATGGCGCAGTTGGGAGAAATGGTG-3' (*Nco*I site  
20 in italics) and 5'-AACAAAGCTTCTTATTGAGGTGCGGGCCGCTTGATGC-3' (*Not*I  
21 site in italics) using standard PCR reaction conditions as described in the Gene Amp kit  
22 (Perkin-Elmer Cetus, CT) and an annealing temperature of 40°C. The template was  
23 plasmid pMON26820. The resulting DNA was digested with *Nco*I and *Not*I and cloned  
24 into a modified pET24d plasmid, designed to add an N-terminal hexa-histidine tag to the  
25 protein. The correct sequence of the PCR product was verified by sequencing, and the  
26 plasmid was transformed into *E.coli* BL21 (DE3), and transformants selected on LB  
27 containing 25 mg/mL kanamycin. The expression strain was grown in LB containing 25  
28 mg/mL kanamycin and induced for 8 hrs at 28°C with 1 mM IPTG. Cells were harvested  
29 and washed in 50 mM Tris/HCl pH 8.5, 150 mM NaCl, and lysed by French Press at  
30 20,000 psi. His-tagged protein was recovered in the soluble fraction of lysed cells and  
31 subsequently purified using Ni-NTA resin as described in the QIAexpressionist manual

1 (Qiagen CA). The partially purified protein was then dialyzed against 25 mM Tris/HCl  
2 pH 7.5 (buffer A) and loaded onto Mono Q HR 10/10 anion-exchange column  
3 (Amersham Pharmacia, NJ) equilibrated with buffer A. The protein was eluted with 25  
4 mM Tris/HCl pH 7.5, 1 M KCl (buffer B) using a linear gradient of 0-100% buffer B run  
5 over 30 min at a flow rate of 4 mL/min using an HPLC system (Shimadzu). Fractions  
6 containing protein were assayed for esterase activity. Esterase active fractions were  
7 pooled, concentrated and dialyzed against 25 mM Tris/HCl pH 7.5 and stored at 4°C.

8 Enzyme Activity Assays

9 Enzyme activity was measured as described previously using *p*-nitrophenyl caprate  
10 (Sigma, MO) as a substrate, dissolved in dimethylsulfoxide (5 mM stock solution) and  
11 diluted in 4% Triton X-100, 1% SDS to a final concentration of 1 mM. For the assay, 20  
12 mL of protein solution was added to a mixture of 25 mL of the 1 mM substrate solution  
13 and 80 mL of 50 mM Tris pH 8.5. The enzyme activity was monitored at 405 nm in 6  
14 sec interval for a period of 10 min. Esterase activity was expressed as DOD min<sup>-1</sup>mg<sup>-1</sup>  
15 protein.

16 Insect Bioassay

17 The protein was also assayed for activity against larvae of *Diabrotica virgifera* (Western  
18 corn rootworm) by overlaying the test sample on an agar diet similar to that described  
19 previously (Marrone et al., *J. Econ. Entom.* 78, 290-3 (1985)). Proteins to be tested were  
20 diluted in 25 mM Tris/HCl pH 7.5 and overlayed on the diet surface. Neonate larvae  
21 were allowed to feed on the diet and mortality and growth stunting were evaluated after 6  
22 days.

23 IgE Binding Epitopes on Patatin

24 A panel of eighty-nine overlapping peptides representing the amino acid sequence of  
25 patatin were synthesized to determine the regions responsible for IgE binding. Each  
26 peptide was 10 amino acids long and consisted of 6 amino acid overlap between the  
27 consecutive peptides. The peptides were evaluated for IgE binding with five different  
28 potato allergic patient sera. Patatin has 3 major epitopes. These major IgE binding  
29 regions represent amino acids 184-193, 188-197, 269-278 and 360-369. Other minor IgE  
30 binding regions represent amino acids 104-113, 138-147 and 316-325. The amino acids

1 essential for IgE binding in each major and minor epitopes were determined by  
2 synthesizing peptides with single amino acid changes at each position by individually  
3 substituting an alanine residue at each non-alanine position in the epitopes. The resulting  
4 alanine substituted peptides were evaluated for IgE binding. Result effective  
5 substitutions were identified by a reduction in IgE binding with respect to the non-  
6 substituted peptide sequence. It was very interesting to note that all the epitopes  
7 contained a Tyr residue and substitution of this Tyr for Ala or Phe eliminated IgE  
8 binding.

9 Enzyme and Bioactivity

10 The Tyr residues identified to be critical for IgE binding in each of the epitopes were  
11 substituted with Phe either individually or in concert using site-directed mutagenesis. All  
12 the variants were expressed in *Pichia pastoris* and assessed for enzyme activity and  
13 insecticidal activity. The variants included Y106F, Y129F, Y185F, Y193F, Y270F,  
14 Y316F, Y362F, Y185F/Y193F, Y185F/Y193F/Y270F/Y316F/Y362F (5-Tyr) and  
15 Y106F/Y129F/Y185F/Y193F/Y270F/Y316F/ Y362F (7-Tyr). All the variants  
16 maintained enzyme activity. The 5-Tyr and 7-Tyr variants were then assessed for  
17 insecticidal activity by overlaying protein (200 ppm final concentration). The proteins  
18 caused significant stunting of the larval growth as measured by the weight of the larvae  
19 after 6 days with the 5-Tyr variant showing higher insecticidal activity compared to the 7-  
20 Tyr and wild type proteins. The 7-Tyr variant was unstable upon long term storage at  
21 4°C and thus was not pursued further.

22 Immunoblotting

23 In order to test if the glycan moiety on patatin was important for binding of IgE, Pat17  
24 was expressed in *E.coli* to produce a nonglycosylated protein and in corn to produce a  
25 plant glycosylated protein. The 5-Tyr variant was also expressed in *E.coli* to assess the  
26 individual contribution of the linear epitopes without the glycan moiety on the protein.  
27 The proteins were tested for binding to IgE using sera from five patients with allergy to  
28 potatoes and sera from one patient with allergies to many things but no allergy to  
29 potatoes. Proteins from both corn and *E.coli* were purified to homogeneity. These  
30 proteins were transferred to PVDF membrane (Millipore, MA) and subsequently probed

1 with sera from patients with and without allergy to potatoes. A His-tagged glyphosate  
2 oxidase control was included in all the studies to verify that the His-tag did not affect the  
3 binding of IgE. Serum obtained from patient HS-07 (no allergy to potatoes) did not bind  
4 Pat17 expressed in *E.coli* but showed good binding to Pat17 from corn and also a protein  
5 at the same molecular weight in potato extract. It is interesting to note that this sera also  
6 showed strong binding to another protein (> 46kDa) in the potato. Sera from patients  
7 HS-01, HS-02, HS-03, HS-05 (allergy to potatoes) shows strong binding to Pat17  
8 expressed in corn, but very weak to no binding to Pat17 produced in *E.coli*. Also, the  
9 sera from patients HS-01, HS-2, HS-03 and HS-05 bound to a protein of similar  
10 molecular weight in the potato extract. Sera from patients HS-01, HS-02 and HS-03 also  
11 showed binding to another protein in potato extract of a lower molecular weight (<  
12 30kDa). Serum obtained from patient HS-06 (allergic to potatoes) showed very strong  
13 binding to wild type patatin expressed in both corn and *E.coli* but weaker binding to the  
14 5-Tyr variant expressed in *E.coli*. Sera from HS-06 also showed very strong binding to a  
15 protein in potato extract with similar molecular weight as patatin. The sera from all the  
16 patients showed no binding to His-tagged glyphosate oxidase indicating that the His-tag  
17 does not bind IgE. These results strongly suggest that the glycan moiety on Pat17 is  
18 responsible for IgE binding in some potato allergic patients and linear epitopes also  
19 contribute to the antigenicity of patatin.

20 Example 20: Alternative nucleic acid and protein sequences

21 For future variations of the patatin protein, sequences showing high similarity to  
22 the sequences disclosed herein could be used in producing deallergenized patatin proteins  
23 and permuteins. For example, a BLAST search (Altschul, S.F. et al., *J. Mol. Biol.* 215:  
24 403-410, 1990) can be performed to identify additional patatin sequences. Sources other  
25 than those disclosed herein can be used to obtain a patatin nucleic acid sequence, and the  
26 encoded patatin protein. Furthermore, subunit sequences from different organisms can be  
27 combined to create a novel patatin sequence incorporating structural, regulatory, and  
28 enzymatic properties from different sources.

1      Example 21: Nucleic acid mutation and hybridization

2           Variations in the nucleic acid sequence encoding a patatin protein may lead to  
3           mutant patatin protein sequences that display equivalent or superior enzymatic  
4           characteristics when compared to the sequences disclosed herein. This invention  
5           accordingly encompasses nucleic acid sequences which are similar to the sequences  
6           disclosed herein, protein sequences which are similar to the sequences disclosed herein,  
7           and the nucleic acid sequences that encode them. Mutations can include deletions,  
8           insertions, truncations, substitutions, fusions, shuffling of subunit sequences, and the like.

9           Mutations to a nucleic acid sequence can be introduced in either a specific or  
10          random manner, both of which are well known to those of skill in the art of molecular  
11          biology. A myriad of site-directed mutagenesis techniques exist, typically using  
12          oligonucleotides to introduce mutations at specific locations in a nucleic acid sequence.  
13          Examples include single strand rescue (Kunkel, T. *Proc. Natl. Acad. Sci. U.S.A.*, 82: 488-  
14          492, 1985), unique site elimination (Deng and Nickloff, *Anal. Biochem.* 200: 81, 1992),  
15          nick protection (Vandeyar, et al. *Gene* 65: 129-133, 1988), and PCR (Costa, et al.  
16          *Methods Mol. Biol.* 57: 31-44, 1996). Random or non-specific mutations can be  
17          generated by chemical agents (for a general review, see Singer and Kusmirek, *Ann. Rev.  
18          Biochem.* 52: 655-693, 1982) such as nitrosoguanidine (Cerda-Olmedo et al., *J. Mol.  
19          Biol.* 33: 705-719, 1968; Guerola, et al. *Nature New Biol.* 230: 122-125, 1971) and 2-  
20          aminopurine (Rogan and Bessman, *J. Bacteriol.* 103: 622-633, 1970), or by biological  
21          methods such as passage through mutator strains (Greener et al. *Mol. Biotechnol.* 7: 189-  
22          195, 1997).

23           Nucleic acid hybridization is a technique well known to those of skill in the art of  
24          DNA manipulation. The hybridization properties of a given pair of nucleic acids is an  
25          indication of their similarity or identity. Mutated nucleic acid sequences can be selected  
26          for their similarity to the disclosed patatin nucleic acid sequences on the basis of their  
27          hybridization to the disclosed sequences. Low stringency conditions can be used to select  
28          sequences with multiple mutations. One may wish to employ conditions such as about  
29          0.15 M to about 0.9 M sodium chloride, at temperatures ranging from about 20°C to  
30          about 55°C. High stringency conditions can be used to select for nucleic acid sequences  
31          with higher degrees of identity to the disclosed sequences. Conditions employed may

1 include about 0.02 M to about 0.15 M sodium chloride, about 0.5% to about 5% casein,  
2 about 0.02% SDS and/or about 0.1% N-laurylsarcosine, about 0.001 M to about 0.03 M  
3 sodium citrate, at temperatures between about 50°C and about 70°C. More preferably,  
4 high stringency conditions are 0.02 M sodium chloride, 0.5% casein, 0.02% SDS, 0.001  
5 M sodium citrate, at a temperature of 50°C.

6 Example 22: Determination of homologous and degenerate nucleic acid sequences

7 Modification and changes can be made in the sequence of the proteins of the  
8 present invention and the nucleic acid segments which encode them and still obtain a  
9 functional molecule that encodes a protein with desirable properties. The following is a  
10 discussion based upon changing the amino acid sequence of a protein to create an  
11 equivalent, or possibly an improved, second-generation molecule. The amino acid  
12 changes can be achieved by changing the codons of the nucleic acid sequence, according  
13 to the codons given in Table 10.

14 Table 10: Codon degeneracies of amino acids

Amino acid	One letter	Three letter	Codons
Alanine	A	Ala	GCA GCC GCG GCT
Cysteine	C	Cys	TGC TGT
Aspartic acid	D	Asp	GAC GAT
Glutamic acid	F	Glu	GAA GAG
Phenylalanine	F	Phe	TTC TTT
Glycine	G	Gly	GGA GGC GGG GGT
Histidine	H	His	CAC CAT
Isoleucine	I	Ile	ATA ATC ATT
Lysine	K	Lys	AAA AAG
Leucine	L	Leu	TTA TTG CTA CTC CTG CTT
Methionine	M	Met	ATG
Asparagine	N	Asn	AAC AAT
Proline	P	Pro	CCA CCC CCG CCT
Glutamine	Q	Gln	CAA CAG
Arginine	R	Arg	AGA AGG CGA CGC CGG CGT
Serine	S	Ser	AGC AGT TCA TCC TCG TCT
Threonine	T	Thr	ACA ACC ACG ACT
Valine	V	Val	GTA GTC GTG GTT
Tryptophan	W	Trp	TGG
Tyrosine	Y	Tyr	TAC TAT

1        Certain amino acids can be substituted for other amino acids in a protein sequence  
2        without appreciable loss of enzymatic activity. It is thus contemplated that various  
3        changes can be made in the peptide sequences of the disclosed protein sequences, or their  
4        corresponding nucleic acid sequences without appreciable loss of the biological activity.

5        In making such changes, the hydropathic index of amino acids can be considered.  
6        The importance of the hydropathic amino acid index in conferring interactive biological  
7        function on a protein is generally understood in the art (Kyte and Doolittle, *J. Mol. Biol.*,  
8        157: 105-132, 1982). It is accepted that the relative hydropathic character of the amino  
9        acid contributes to the secondary structure of the resultant protein, which in turn defines  
10      the interaction of the protein with other molecules, for example, enzymes, substrates,  
11      receptors, DNA, antibodies, antigens, and the like.

12      Each amino acid has been assigned a hydropathic index on the basis of their  
13      hydrophobicity and charge characteristics. These are: isoleucine (+4.5); valine (+4.2);  
14      leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine  
15      (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3);  
16      proline (-1.6); histidine (-3.2); glutamate/glutamine/aspartate/asparagine (-3.5); lysine (-  
17      3.9); and arginine (-4.5).

18      It is known in the art that certain amino acids can be substituted by other amino  
19      acids having a similar hydropathic index or score and still result in a protein with similar  
20      biological activity, i.e., still obtain a biologically functional protein. In making such  
21      changes, the substitution of amino acids whose hydropathic indices are within  $\pm 2$  is  
22      preferred, those within  $\pm 1$  are more preferred, and those within  $\pm 0.5$  are most preferred.

23      It is also understood in the art that the substitution of like amino acids can be  
24      made effectively on the basis of hydrophilicity. U.S. Patent No. 4,554,101 (Hopp, T.P.,  
25      issued November 19, 1985) states that the greatest local average hydrophilicity of a  
26      protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a  
27      biological property of the protein. The following hydrophilicity values have been  
28      assigned to amino acids: arginine/lysine (+3.0); aspartate/glutamate (+3.0  $\pm 1$ ); serine  
29      (+0.3); asparagine/glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5  $\pm 1$ );  
30      alanine/histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5);  
31      leucine/isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); and tryptophan (-3.4).

1 It is understood that an amino acid can be substituted by another amino acid  
2 having a similar hydrophilicity score and still result in a protein with similar biological  
3 activity, i.e., still obtain a biologically functional protein. In making such changes, the  
4 substitution of amino acids whose hydropathic indices are within  $\pm 2$  is preferred, those  
5 within  $\pm 1$  are more preferred, and those within  $\pm 0.5$  are most preferred.

6 As outlined above, amino acid substitutions are therefore based on the relative  
7 similarity of the amino acid side-chain constituents, for example, their hydrophobicity,  
8 hydrophilicity, charge, size, and the like. Exemplary substitutions which take various of  
9 the foregoing characteristics into consideration are well known to those of skill in the art  
10 and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine  
11 and asparagine; and valine, leucine, and isoleucine. Changes which are not expected to  
12 be advantageous may also be used if these resulted in functional patatin proteins.

13 Example 23: Production of patatin proteins and permuteins in plants

14 Plant Vectors

15 In plants, transformation vectors capable of introducing nucleic acid sequences  
16 encoding patatin proteins and permuteins are easily designed, and generally contain one  
17 or more nucleic acid coding sequences of interest under the transcriptional control of 5'  
18 and 3' regulatory sequences. Such vectors generally comprise, operatively linked in  
19 sequence in the 5' to 3' direction, a promoter sequence that directs the transcription of a  
20 downstream heterologous structural nucleic acid sequence in a plant; optionally, a 5' non-  
21 translated leader sequence; a nucleic acid sequence that encodes a protein of interest; and  
22 a 3' non-translated region that encodes a polyadenylation signal which functions in plant  
23 cells to cause the termination of transcription and the addition of polyadenylate  
24 nucleotides to the 3' end of the mRNA encoding the protein. Plant transformation  
25 vectors also generally contain a selectable marker. Typical 5'-3' regulatory sequences  
26 include a transcription initiation start site, a ribosome binding site, an RNA processing  
27 signal, a transcription termination site, and/or a polyadenylation signal. Vectors for plant  
28 transformation have been reviewed in Rodriguez et al. (Vectors: A Survey of Molecular  
29 Cloning Vectors and Their Uses, Butterworths, Boston., 1988), Glick et al. (Methods in  
30 Plant Molecular Biology and Biotechnology, CRC Press, Boca Raton, Fla., 1993), and

1 Croy (Plant Molecular Biology Labfax, Hames and Rickwood (Eds.), BIOS Scientific  
2 Publishers Limited, Oxford, UK., 1993).

3 Plant Promoters

4 Plant promoter sequences can be constitutive or inducible, environmentally- or  
5 developmentally-regulated, or cell- or tissue-specific. Often-used constitutive promoters  
6 include the CaMV 35S promoter (Odell, J.T. et al., *Nature* 313: 810-812, 1985), the  
7 enhanced CaMV 35S promoter, the Figwort Mosaic Virus (FMV) promoter (Richins et  
8 al., *Nucleic Acids Res.* 20: 8451-8466, 1987), the mannopine synthase (*mas*) promoter,  
9 the nopaline synthase (*nos*) promoter, and the octopine synthase (*ocs*) promoter. Useful  
10 inducible promoters include promoters induced by salicylic acid or polyacrylic acids (PR-  
11 1, Williams , S. W. et al, *Biotechnology* 10: 540-543, 1992), induced by application of  
12 safeners (substituted benzenesulfonamide herbicides, Hershey, H.P. and Stoner, T.D.,  
13 *Plant Mol. Biol.* 17: 679-690, 1991), heat-shock promoters (Ou-Lee et al., *Proc. Natl.  
14 Acad. Sci. U.S.A.* 83: 6815-6819, 1986; Ainley et al., *Plant Mol. Biol.* 14: 949-967,  
15 1990), a nitrate-inducible promoter derived from the spinach nitrite reductase gene (Back  
16 et al., *Plant Mol. Biol.* 17: 9-18, 1991), hormone-inducible promoters (Yamaguchi-  
17 Shinozaki, K. et al., *Plant Mol. Biol.* 15: 905-912, 1990; Kares et al., *Plant Mol. Biol.* 15:  
18 225-236, 1990), and light-inducible promoters associated with the small subunit of RuBP  
19 carboxylase and LHCP gene families (Kuhlemeier et al., *Plant Cell* 1: 471, 1989;  
20 Feinbaum, R.L. et al., *Mol. Gen. Genet.* 226: 449-456, 1991; Weisshaar, B. et al., *EMBO  
21 J.* 10: 1777-1786, 1991; Lam, E. and Chua, N.H., *J. Biol. Chem.* 266: 17131-17135,  
22 1990; Castresana, C. et al., *EMBO J.* 7: 1929-1936, 1988; Schulze-Lefert et al., *EMBO J.*  
23 8: 651, 1989). Examples of useful tissue-specific, developmentally-regulated promoters  
24 include the  $\beta$ -conglycinin 7S promoter (Doyle, J.J. et al., *J. Biol. Chem.* 261: 9228-9238,  
25 1986; Slighton and Beachy, *Planta* 172: 356-363, 1987), and seed-specific promoters  
26 (Knutzon, D.S. et al., *Proc. Natl. Acad. Sci. U.S.A.* 89: 2624-2628, 1992; Bustos, M.M. et  
27 al., *EMBO J.* 10: 1469-1479, 1991; Lam and Chua, *Science* 248: 471, 1991; Stayton et  
28 al., *Aust. J. Plant. Physiol.* 18: 507, 1991). Plant functional promoters useful for  
29 preferential expression in seed plastids include those from plant storage protein genes and  
30 from genes involved in fatty acid biosynthesis in oilseeds. Examples of such promoters

1 include the 5' regulatory regions from such genes as napin (Kridl et al., *Seed Sci. Res.* 1: 209-219, 1991), phaseolin, zein, soybean trypsin inhibitor, ACP, stearoyl-ACP 2 desaturase, and oleosin. Seed-specific gene regulation is discussed in EP 0 255 378. 3 Promoter hybrids can also be constructed to enhance transcriptional activity (Comai, L. 4 and Moran, P.M., U.S. Patent No. 5,106,739, issued April 21, 1992), or to combine 5 desired transcriptional activity and tissue specificity.

6

7 Plant transformation and regeneration

8 A variety of different methods can be employed to introduce such vectors into 9 plant protoplasts, cells, callus tissue, leaf discs, meristems, etcetera, to generate 10 transgenic plants, including *Agrobacterium*-mediated transformation, particle gun 11 delivery, microinjection, electroporation, polyethylene glycol mediated protoplast 12 transformation, liposome-mediated transformation, etcetera (reviewed in Potrykus, I. 13 *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 42: 205-225, 1991). In general, transgenic 14 plants comprising cells containing and expressing DNAs encoding patatin proteins and 15 permuteins can be produced by transforming plant cells with a DNA construct as 16 described above via any of the foregoing methods; selecting plant cells that have been 17 transformed on a selective medium; regenerating plant cells that have been transformed 18 to produce differentiated plants; and selecting a transformed plant which expresses the 19 protein-encoding nucleotide sequence.

20 Specific methods for transforming a wide variety of dicots and obtaining 21 transgenic plants are well documented in the literature (Gasser and Fraley, *Science* 244: 22 1293-1299, 1989; Fisk and Dandekar, *Scientia Horticulturae* 55: 5-36, 1993; Christou, 23 *Agro Food Industry Hi Tech*, p.17, 1994; and the references cited therein).

24 Successful transformation and plant regeneration have been reported in the 25 monocots as follows: asparagus (*Asparagus officinalis*; Bytebier et al., *Proc. Natl. Acad. 26 Sci. U.S.A.* 84: 5345-5349, 1987); barley (*Hordeum vulgare*; Wan and Lemaux, *Plant 27 Physiol.* 104: 37-48, 1994); maize (*Zea mays*; Rhodes, C.A. et al., *Science* 240: 204-207, 28 1988; Gordon-Kamm et al., *Plant Cell* 2: 603-618, 1990; Fromm, M.E. et al., 29 *Bio/Technology* 8: 833-839, 1990; Koziel et al., *Bio/Technology* 11: 194-200, 1993); oats 30 (*Avena sativa*; Somers et al., *Bio/Technology* 10: 1589-1594, 1992); orchardgrass

1 (*Dactylis glomerata*; Horn et al., *Plant Cell Rep.* 7: 469-472, 1988); rice (*Oryza sativa*,  
2 including indica and japonica varieties; Toriyama et al., *Bio/Technology* 6: 10, 1988;  
3 Zhang et al., *Plant Cell Rep.* 7: 379-384, 1988; Luo and Wu, *Plant Mol. Biol. Rep.* 6:  
4 165-174, 1988; Zhang and Wu, *Theor. Appl. Genet.* 76: 835-840, 1988; Christou et al.,  
5 *Bio/Technology* 9: 957-962, 1991); rye (*Secale cereale*; De la Pena et al., *Nature* 325:  
6 274-276, 1987); sorghum (*Sorghum bicolor*; Casas, A.M. et al., *Proc. Natl. Acad. Sci.*  
7 U.S.A. 90: 11212-11216, 1993); sugar cane (*Saccharum* spp.; Bower and Birch, *Plant J.*  
8 2: 409-416, 1992); tall fescue (*Festuca arundinacea*; Wang, Z.Y. et al., *Bio/Technology*  
9 10: 691-696, 1992); turfgrass (*Agrostis palustris*; Zhong et al., *Plant Cell Rep.* 13: 1-6,  
10 1993); wheat (*Triticum aestivum*; Vasil et al., *Bio/Technology* 10: 667-674, 1992; Weeks,  
11 T. et al., *Plant Physiol.* 102: 1077-1084, 1993; Becker et al., *Plant J.* 5: 299-307, 1994),  
12 and alfalfa (Masoud, S.A. et al., *Transgen. Res.* 5: 313, 1996); *Brassica* (canola/oilseed  
13 rape) (Fry, J. *Plant Cell Rep.* 6: 321-325, 1987); and soybean (Hinchee, M. *Bio/Technol.*  
14 6: 915-922, 1988).

15 All of the compositions and/or methods disclosed and claimed herein can be made  
16 and executed without undue experimentation in light of the present disclosure. While the  
17 compositions and methods of this invention have been described in terms of preferred  
18 embodiments, it will be apparent to those of skill in the art that variations can be applied  
19 to the compositions and/or methods and in the steps or in the sequence of steps of the  
20 methods described herein without departing from the concept, spirit and scope of the  
21 invention. More specifically, it will be apparent that certain agents which are both  
22 chemically and physiologically related can be substituted for the agents described herein  
23 while the same or similar results would be achieved. All such similar substitutes and  
24 modifications apparent to those skilled in the art are deemed to be within the spirit, scope  
25 and concept of the invention.

26

27

## SEQUENCE LISTING

1 acctatgagg aagctctaaa gaggttgca aaattgctct ctgatagga acaaactccga 1140  
 2  
 3 gcaaaacaaag ctttttat 1158  
 4  
 5 <210> 2  
 6 <211> 386  
 7 <212> PRT  
 8 <213> Solanum Tuberosum  
 9  
 10 <400> 2  
 11  
 12  
 13 Met Ala Thr Thr Lys Ser Phe Leu Ile Leu Ile Phe Met Ile Leu Ala  
 14 1 5 10 15  
 15  
 16 Thr Thr Ser Ser Thr Phe Ala Gln Leu Gly Glu Met Val Thr Val Leu  
 17 20 25 30  
 18  
 19 Ser Ile Asp Gly Gly Ile Arg Gly Ile Ile Pro Ala Thr Ile Leu  
 20 35 40 45  
 21  
 22 Glu Phe Leu Glu Gly Gln Leu Gln Glu Met Asp Asn Asn Ala Asp Ala  
 23 50 55 60  
 24  
 25 Arg Leu Ala Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser Thr Gly Gly  
 26 65 70 75 80  
 27  
 28 Leu Leu Thr Ala Met Ile Ser Thr Pro Asn Glu Asn Asn Arg Pro Phe  
 29 85 90 95  
 30  
 31 Ala Ala Ala Lys Glu Ile Val Pro Phe Tyr Phe Glu His Gly Pro Gln  
 32 100 105 110  
 33  
 34 Ile Phe Asn Pro Ser Gly Gln Ile Leu Gly Pro Lys Tyr Asp Gly Lys  
 35 115 120 125  
 36  
 37 Tyr Leu Met Gln Val Leu Gln Glu Lys Leu Gly Glu Thr Arg Val His  
 38 130 135 140  
 39  
 40 Gln Ala Leu Thr Glu Val Val Ile Ser Ser Phe Asp Ile Lys Thr Asn  
 41 145 150 155 160  
 42  
 43 Lys Pro Val Ile Phe Thr Lys Ser Asn Leu Ala Asn Ser Pro Glu Leu  
 44 165 170 175  
 45  
 46 Asp Ala Lys Met Tyr Asp Ile Ser Tyr Ser Thr Ala Ala Ala Pro Thr  
 47 180 185 190  
 48  
 49 Tyr Phe Pro Pro His Tyr Phe Val Thr Asn Thr Ser Asn Gly Asp Glu  
 50 195 200 205  
 51  
 52 Tyr Glu Phe Asn Leu Val Asp Gly Ala Val Ala Thr Val Ala Asp Pro  
 53 210 215 220  
 54  
 55 Ala Leu Leu Ser Ile Ser Val Ala Thr Arg Leu Ala Gln Lys Asp Pro  
 56 225 230 235 240  
 57

1 Ala Phe Ala Ser Ile Arg Ser Leu Asn Tyr Lys Lys Met Leu Leu Leu  
 2 245 250 255  
 3  
 4 Ser Leu Gly Thr Gly Thr Thr Ser Glu Phe Asp Lys Thr Tyr Thr Ala  
 5 260 265 270  
 6  
 7 Lys Glu Ala Ala Thr Trp Thr Ala Val His Trp Met Leu Val Ile Gln  
 8 275 280 285  
 9  
 10 Lys Met Thr Asp Ala Ala Ser Ser Tyr Met Thr Asp Tyr Tyr Leu Ser  
 11 290 295 300  
 12  
 13 Thr Ala Phe Gln Ala Leu Asp Ser Lys Asn Asn Tyr Leu Arg Val Gln  
 14 305 310 315 320  
 15  
 16 Glu Asn Ala Leu Thr Gly Thr Thr Glu Met Asp Asp Ala Ser Glu  
 17 325 330 335  
 18  
 19 Ala Asn Met Glu Leu Leu Val Gln Val Gly Glu Asn Leu Leu Lys Lys  
 20 340 345 350  
 21  
 22 Pro Val Ser Glu Asp Asn Pro Glu Thr Tyr Glu Glu Ala Leu Lys Arg  
 23 355 360 365  
 24  
 25 Phe Ala Lys Leu Leu Ser Asp Arg Lys Lys Leu Arg Ala Asn Lys Ala  
 26 370 375 380  
 27  
 28 Ser Tyr  
 29 385  
 30  
 31 <210> 3  
 32 <211> 54  
 33 <212> DNA  
 34 <213> Synthetic  
 35  
 36 <400> 3  
 37 ggagctcgag aaaagagagg ctgaagctca gttgggagaa atggtgactg ttct 54  
 38  
 39 <210> 4  
 40 <211> 29  
 41 <212> DNA  
 42 <213> Synthetic  
 43  
 44 <400> 4  
 45 ggtctagagg aattctcatt aataagaag 29  
 46  
 47 <210> 5  
 48 <211> 1138  
 49 <212> DNA  
 50 <213> Synthetic  
 51  
 52 <400> 5  
 53 ggagctcgag aaaagagagg ctgaagctca gttgggagaa atggtgactg ttcttagtat 60  
 54  
 55 tggatggatgg ggaatttagag ggatcattcc ggctaccatt ctcaatttc ttgaaggaca 120  
 56  
 57 acttcaggaa atggacaata atgcagatgc aagacttgca gattactttg atgttaattgg 180



1 Val Leu Ser Ile Asp Gly Gly Gly Ile Arg Gly Ile Ile Pro Ala Thr  
2 100 105 110  
3  
4 Ile Leu Glu Phe Leu Glu Gly Gln Leu Gln Glu Met Asp Asn Asn Ala  
5 115 120 125  
6  
7 Asp Ala Arg Leu Ala Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser Thr  
8 130 135 140  
9  
10 Gly Gly Leu Leu Thr Ala Met Ile Ser Thr Pro Asn Glu Asn Asn Arg  
11 145 150 155 160  
12  
13 Pro Phe Ala Ala Ala Lys Glu Ile Val Pro Phe Tyr Phe Glu His Gly  
14 165 170 175  
15  
16 Pro Gln Ile Phe Asn Pro Ser Gly Gln Ile Leu Gly Pro Lys Tyr Asp  
17 180 185 190  
18  
19 Gly Lys Tyr Leu Met Gln Val Leu Gln Glu Lys Leu Gly Glu Thr Arg  
20 195 200 205  
21  
22 Val His Gln Ala Leu Thr Glu Val Val Ile Ser Ser Phe Asp Ile Lys  
23 210 215 220  
24  
25 Thr Asn Lys Pro Val Ile Phe Thr Lys Ser Asn Leu Ala Asn Ser Pro  
26 225 230 235 240  
27  
28 Glu Leu Asp Ala Lys Met Tyr Asp Ile Ser Tyr Ser Thr Ala Ala Ala  
29 245 250 255  
30  
31 Pro Thr Tyr Phe Pro Pro His Tyr Phe Val Thr Asn Thr Ser Asn Gly  
32 260 265 270  
33  
34 Asp Glu Tyr Glu Phe Asn Leu Val Asp Gly Ala Val Ala Thr Val Ala  
35 275 280 285  
36  
37 Asp Pro Ala Leu Leu Ser Ile Ser Val Ala Thr Arg Leu Ala Gln Lys  
38 290 295 300  
39  
40 Asp Pro Ala Phe Ala Ser Ile Arg Ser Leu Asn Tyr Lys Lys Met Leu  
41 305 310 315 320  
42  
43 Leu Leu Ser Leu Gly Thr Gly Thr Ser Glu Phe Asp Lys Thr Tyr  
44 325 330 335  
45  
46 Thr Ala Lys Glu Ala Ala Thr Trp Thr Ala Val His Trp Met Leu Val  
47 340 345 350  
48  
49 Ile Gln Lys Met Thr Asp Ala Ala Ser Ser Tyr Met Thr Asp Tyr Tyr  
50 355 360 365  
51  
52 Leu Ser Thr Ala Phe Gln Ala Leu Asp Ser Lys Asn Asn Tyr Leu Arg  
53 370 375 380  
54  
55 Val Gln Glu Asn Ala Leu Thr Gly Thr Thr Glu Met Asp Asp Ala  
56 385 390 395 400  
57

1 Ser Glu Ala Asn Met Glu Leu Leu Val Gln Val Gly Glu Asn Leu Leu  
 2 405 410 415  
 3  
 4 Lys Lys Pro Val Ser Glu Asp Asn Pro Glu Thr Tyr Glu Glu Ala Leu  
 5 420 425 430  
 6  
 7 Lys Arg Phe Ala Lys Leu Leu Ser Asp Arg Lys Lys Leu Arg Ala Asn  
 8 435 440 445  
 9  
 10 Lys Ala Ser Tyr  
 11 450  
 12  
 13  
 14 <210> 7  
 15 <211> 367  
 16 <212> PRT  
 17 <213> Synthetic  
 18  
 19 <400> 7  
 20  
 21 Glu Ala Glu Ala Gln Leu Gly Glu Met Val Thr Val Leu Ser Ile Asp  
 22 1 5 10 15  
 23  
 24 Gly Gly Gly Ile Arg Gly Ile Ile Pro Ala Thr Ile Leu Glu Phe Leu  
 25 20 25 30  
 26  
 27 Glu Gly Gln Leu Gln Glu Met Asp Asn Asn Ala Asp Ala Arg Leu Ala  
 28 35 40 45  
 29  
 30 Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser Thr Gly Gly Leu Leu Thr  
 31 50 55 60  
 32  
 33 Ala Met Ile Ser Thr Pro Asn Glu Asn Asn Arg Pro Phe Ala Ala Ala  
 34 65 70 75 80  
 35  
 36 Lys Glu Ile Val Pro Phe Tyr Phe Glu His Gly Pro Gln Ile Phe Asn  
 37 85 90 95  
 38  
 39 Pro Ser Gly Gln Ile Leu Gly Pro Lys Tyr Asp Gly Lys Tyr Leu Met  
 40 100 105 110  
 41  
 42 Gln Val Leu Gln Glu Lys Leu Gly Glu Thr Arg Val His Gln Ala Leu  
 43 115 120 125  
 44  
 45 Thr Glu Val Val Ile Ser Ser Phe Asp Ile Lys Thr Asn Lys Pro Val  
 46 130 135 140  
 47  
 48 Ile Phe Thr Lys Ser Asn Leu Ala Asn Ser Pro Glu Leu Asp Ala Lys  
 49 145 150 155 160  
 50  
 51 Met Tyr Asp Ile Ser Tyr Ser Thr Ala Ala Ala Pro Thr Tyr Phe Pro  
 52 165 170 175  
 53  
 54 Pro His Tyr Phe Val Thr Asn Thr Ser Asn Gly Asp Glu Tyr Glu Phe  
 55 180 185 190  
 56  
 57 Asn Leu Val Asp Gly Ala Val Ala Thr Val Ala Asp Pro Ala Leu Leu

1	195	200	205	
2	Ser Ile Ser Val Ala Thr Arg Leu Ala Gln Lys Asp Pro Ala Phe Ala			
3	210	215	220	
4	Ser Ile Arg Ser Leu Asn Tyr Lys Lys Met Leu Leu Leu Ser Leu Gly			
5	225	230	235	240
6	Thr Gly Thr Thr Ser Glu Phe Asp Lys Thr Tyr Thr Ala Lys Glu Ala			
7	245	250	255	
8	Ala Thr Trp Thr Ala Val His Trp Met Leu Val Ile Gln Lys Met Thr			
9	260	265	270	
10	Asp Ala Ala Ser Ser Tyr Met Thr Asp Tyr Tyr Leu Ser Thr Ala Phe			
11	275	280	285	
12	Gln Ala Leu Asp Ser Lys Asn Asn Tyr Leu Arg Val Gln Glu Asn Ala			
13	290	295	300	
14	Leu Thr Gly Thr Thr Glu Met Asp Asp Ala Ser Glu Ala Asn Met			
15	305	310	315	320
16	Glu Leu Leu Val Gln Val Gly Glu Asn Leu Leu Lys Lys Pro Val Ser			
17	325	330	335	
18	Glu Asp Asn Pro Glu Thr Tyr Glu Glu Ala Leu Lys Arg Phe Ala Lys			
19	340	345	350	
20	Leu Leu Ser Asp Arg Lys Lys Leu Arg Ala Asn Lys Ala Ser Tyr			
21	355	360	365	
22	<210> 8			
23	<211> 24			
24	<212> DNA			
25	<213> Synthetic			
26	<400> 8			
27	atgttcgaag aaaaaaggta caat			24
28	<210> 9			
29	<211> 24			
30	<212> DNA			
31	<213> Synthetic			
32	<400> 9			
33	ttgcataaga aattttccat cata			24
34	<210> 10			
35	<211> 24			
36	<212> DNA			
37	<213> Synthetic			
38	<400> 10			
39	tgttgtggaa aaacttatgt cata			24
40	<210> 11			

```

1  <211> 24
2  <212> DNA
3  <213> Synthetic
4
5  <400> 11
6  cggaggaaaa aatgttggag ctgc
7
8  <210> 12
9  <211> 51
10 <212> DNA
11 <213> Synthetic
12
13 <400> 12
14 atgcggagga aaaaatgtg gagctgctgc tgtggaaaaa cttatgtcat a
15
16 <210> 13
17 <211> 24
18 <212> DNA
19 <213> Synthetic
20
21 <400> 13
22 ttttgcgtta aatgttttat caaa
23
24 <210> 14
25 <211> 24
26 <212> DNA
27 <213> Synthetic
28
29 <400> 14
30 aaccctgagg aaattgtttt ttga
31
32 <210> 15
33 <211> 24
34 <212> DNA
35 <213> Synthetic
36
37 <400> 15
38 agcttcctca aaggtttcag gatt
39
40 <210> 16
41 <211> 10
42 <212> PRT
43 <213> Synthetic
44
45 <400> 16
46
47 Gln Leu Gly Glu Met Val Thr Val Leu Ser
48 1 5 10
49
50 <210> 17
51 <211> 10
52 <212> PRT
53 <213> Synthetic
54
55 <400> 17
56
57 Met Val Thr Val Leu Ser Ile Asp Gly

```

1 1 5 10  
2 <210> 18  
3 <211> 10  
4 <212> PRT  
5 <213> Synthetic  
6  
7 <400> 18  
8  
9 Leu Ser Ile Asp Gly Gly Gly Ile Arg Gly  
10 1 5 10  
11  
12 <210> 19  
13 <211> 10  
14 <212> PRT  
15 <213> Synthetic  
16  
17 <400> 19  
18  
19 Gly Gly Gly Ile Arg Gly Ile Ile Pro Ala  
20 1 5 10  
21  
22 <210> 20  
23 <211> 10  
24 <212> PRT  
25 <213> Synthetic  
26  
27 <400> 20  
28  
29 Arg Gly Ile Ile Pro Ala Thr Ile Leu Glu  
30 1 5 10  
31  
32 <210> 21  
33 <211> 10  
34 <212> PRT  
35 <213> Synthetic  
36  
37 <400> 21  
38  
39 Pro Ala Thr Ile Leu Glu Phe Leu Glu Gly  
40 1 5 10  
41  
42 <210> 22  
43 <211> 10  
44 <212> PRT  
45 <213> Synthetic  
46  
47 <400> 22  
48  
49 Leu Glu Phe Leu Glu Gly Gln Leu Gln Glu  
50 1 5 10  
51  
52 <210> 23  
53 <211> 10  
54 <212> PRT  
55 <213> Synthetic  
56  
57 <400> 23

1  
2 Glu Gly Gln Leu Gln Glu Met Asp Asn Asn  
3 1 5 10  
4  
5 <210> 24  
6 <211> 10  
7 <212> PRT  
8 <213> Synthetic  
9  
10 <400> 24  
11  
12 Gln Glu Met Asp Asn Asn Ala Asp Ala Arg  
13 1 5 10  
14  
15 <210> 25  
16 <211> 10  
17 <212> PRT  
18 <213> Synthetic  
19  
20 <400> 25  
21  
22 Asn Asn Ala Asp Ala Arg Leu Ala Asp Tyr  
23 1 5 10  
24  
25 <210> 26  
26 <211> 10  
27 <212> PRT  
28 <213> Synthetic  
29  
30 <400> 26  
31  
32 Ala Arg Leu Ala Asp Tyr Phe Asp Val Ile  
33 1 5 10  
34  
35 <210> 27  
36 <211> 10  
37 <212> PRT  
38 <213> Synthetic  
39  
40 <400> 27  
41  
42 Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser  
43 1 5 10  
44  
45 <210> 28  
46 <211> 10  
47 <212> PRT  
48 <213> Synthetic  
49  
50 <400> 28  
51  
52 Val Ile Gly Gly Thr Ser Thr Gly Gly Leu  
53 1 5 10  
54  
55 <210> 29  
56 <211> 10  
57 <212> PRT

1 <213> Synthetic  
2  
3 <400> 29  
4  
5 Thr Ser Thr Gly Gly Leu Leu Thr Ala Met  
6 1 5 10  
7  
8 <210> 30  
9 <211> 10  
10 <212> PRT  
11 <213> Synthetic  
12  
13 <400> 30  
14  
15 Gly Leu Leu Thr Ala Met Ile Ser Thr Pro  
16 1 5 10  
17  
18 <210> 31  
19 <211> 10  
20 <212> PRT  
21 <213> Synthetic  
22  
23 <400> 31  
24  
25 Ala Met Ile Ser Thr Pro Asn Glu Asn Asn  
26 1 5 10  
27  
28 <210> 32  
29 <211> 10  
30 <212> PRT  
31 <213> Synthetic  
32  
33 <400> 32  
34  
35 Thr Pro Asn Glu Asn Asn Arg Pro Phe Ala  
36 1 5 10  
37  
38 <210> 33  
39 <211> 10  
40 <212> PRT  
41 <213> Synthetic  
42  
43 <400> 33  
44  
45 Asn Asn Arg Pro Phe Ala Ala Ala Lys Glu  
46 1 5 10  
47  
48 <210> 34  
49 <211> 10  
50 <212> PRT  
51 <213> Synthetic  
52  
53 <400> 34  
54  
55 Phe Ala Ala Ala Lys Glu Ile Val Pro Phe  
56 1 5 10  
57

1 <210> 35  
2 <211> 10  
3 <212> PRT  
4 <213> Synthetic  
5 <400> 35  
6  
7 Lys Glu Ile Val Pro Phe Tyr Phe Glu His  
8 1 5 10  
9  
10 <210> 36  
11 <211> 10  
12 <212> PRT  
13 <213> Synthetic  
14 <400> 36  
15  
16 <210> 37  
17 <211> 10  
18 <212> PRT  
19 <213> Synthetic  
20 <400> 37  
21  
22 <210> 38  
23 <211> 10  
24 <212> PRT  
25 <213> Synthetic  
26 <400> 38  
27  
28 Glu His Gly Pro Gln Ile Phe Asn Pro Ser  
29 1 5 10  
30  
31 <210> 39  
32 <211> 10  
33 <212> PRT  
34 <213> Synthetic  
35  
36 <400> 39  
37  
38 Gln Ile Phe Asn Pro Ser Gly Gln Ile Leu  
39 1 5 10  
40  
41 <210> 40  
42 <211> 10  
43 <212> PRT  
44 <213> Synthetic  
45  
46 <400> 40  
47  
48 Pro Ser Gly Gln Ile Leu Gly Pro Lys Tyr  
49 1 5 10  
50  
51 <210> 40  
52 <211> 10  
53 <212> PRT  
54 <213> Synthetic  
55  
56 <400> 40  
57

1 Ile Leu Gly Pro Lys Tyr Asp Gly Lys Tyr  
2 1 5 10  
3  
4 <210> 41  
5 <211> 10  
6 <212> PRT  
7 <213> Synthetic  
8  
9 <400> 41  
10  
11 Lys Tyr Asp Gly Lys Tyr Leu Met Gln Val  
12 1 5 10  
13  
14 <210> 42  
15 <211> 10  
16 <212> PRT  
17 <213> Synthetic  
18  
19 <400> 42  
20  
21 Lys Tyr Leu Met Gln Val Leu Gln Glu Lys  
22 1 5 10  
23  
24 <210> 43  
25 <211> 10  
26 <212> PRT  
27 <213> Synthetic  
28  
29 <400> 43  
30  
31 Gln Val Leu Gln Glu Lys Leu Gly Glu Thr  
32 1 5 10  
33  
34 <210> 44  
35 <211> 10  
36 <212> PRT  
37 <213> Synthetic  
38  
39 <400> 44  
40  
41 Glu Lys Leu Gly Glu Thr Arg Val His Gln  
42 1 5 10  
43  
44 <210> 45  
45 <211> 10  
46 <212> PRT  
47 <213> Synthetic  
48  
49 <400> 45  
50  
51 Glu Thr Arg Val His Gln Ala Leu Thr Glu  
52 1 5 10  
53  
54 <210> 46  
55 <211> 10  
56 <212> PRT  
57 <213> Synthetic

1  
2 <400> 46  
3  
4 His Gln Ala Leu Thr Glu Val Val Ile Ser  
5 1 5 10  
6  
7 <210> 47  
8 <211> 10  
9 <212> PRT  
10 <213> Synthetic  
11  
12 <400> 47  
13  
14 Thr Glu Val Val Ile Ser Ser Phe Asp Ile  
15 1 5 10  
16  
17 <210> 48  
18 <211> 10  
19 <212> PRT  
20 <213> Synthetic  
21  
22 <400> 48  
23  
24 Ile Ser Ser Phe Asp Ile Lys Thr Asn Lys  
25 1 5 10  
26  
27 <210> 49  
28 <211> 10  
29 <212> PRT  
30 <213> Synthetic  
31  
32 <400> 49  
33  
34 Asp Ile Lys Thr Asn Lys Pro Val Ile Phe  
35 1 5 10  
36  
37 <210> 50  
38 <211> 10  
39 <212> PRT  
40 <213> Synthetic  
41  
42 <400> 50  
43  
44 Asn Lys Pro Val Ile Phe Thr Lys Ser Asn  
45 1 5 10  
46  
47 <210> 51  
48 <211> 10  
49 <212> PRT  
50 <213> Synthetic  
51  
52 <400> 51  
53  
54 Ile Phe Thr Lys Ser Asn Leu Ala Asn Ser  
55 1 5 10  
56  
57 <210> 52

1 <211> 10  
2 <212> PRT  
3 <213> Synthetic  
4  
5 <400> 52  
6  
7 Ser Asn Leu Ala Asn Ser Pro Glu Leu Asp  
8 1 5 10  
9  
10 <210> 53  
11 <211> 10  
12 <212> PRT  
13 <213> Synthetic  
14  
15 <400> 53  
16  
17 Asn Ser Pro Glu Leu Asp Ala Lys Met Tyr  
18 1 5 10  
19  
20 <210> 54  
21 <211> 10  
22 <212> PRT  
23 <213> Synthetic  
24  
25 <400> 54  
26  
27 Leu Asp Ala Lys Met Tyr Asp Ile Ser Tyr  
28 1 5 10  
29  
30 <210> 55  
31 <211> 10  
32 <212> PRT  
33 <213> Synthetic  
34  
35 <400> 55  
36  
37 Met Tyr Asp Ile Ser Tyr Ser Thr Ala Ala  
38 1 5 10  
39  
40 <210> 56  
41 <211> 10  
42 <212> PRT  
43 <213> Synthetic  
44  
45 <400> 56  
46  
47 Ser Tyr Ser Thr Ala Ala Ala Pro Thr Tyr  
48 1 5 10  
49  
50 <210> 57  
51 <211> 10  
52 <212> PRT  
53 <213> Synthetic  
54  
55 <400> 57  
56  
57 Ala Ala Ala Pro Thr Tyr Phe Pro Pro His

1 1 5 10  
2  
3 <210> 58  
4 <211> 10  
5 <212> PRT  
6 <213> Synthetic  
7  
8 <400> 58  
9  
10 Thr Tyr Phe Pro Pro His Tyr Phe Val Thr  
11 1 5 10  
12  
13 <210> 59  
14 <211> 10  
15 <212> PRT  
16 <213> Synthetic  
17  
18 <400> 59  
19  
20 Pro His Tyr Phe Val Thr Asn Thr Ser Asn  
21 1 5 10  
22  
23 <210> 60  
24 <211> 10  
25 <212> PRT  
26 <213> Synthetic  
27  
28 <400> 60  
29  
30 Val Thr Asn Thr Ser Asn Gly Asp Glu Tyr  
31 1 5 10  
32  
33 <210> 61  
34 <211> 10  
35 <212> PRT  
36 <213> Synthetic  
37  
38 <400> 61  
39  
40 Ser Asn Gly Asp Glu Tyr Glu Phe Asn Leu  
41 1 5 10  
42  
43 <210> 62  
44 <211> 10  
45 <212> PRT  
46 <213> Synthetic  
47  
48 <400> 62  
49  
50 Glu Tyr Glu Phe Asn Leu Val Asp Gly Ala  
51 1 5 10  
52  
53 <210> 63  
54 <211> 10  
55 <212> PRT  
56 <213> Synthetic  
57

1 <400> 63  
2  
3 Asn Leu Val Asp Gly Ala Val Ala Thr Val  
4 1 5 10  
5  
6 <210> 64  
7 <211> 10  
8 <212> PRT  
9 <213> Synthetic  
10  
11 <400> 64  
12  
13 Gly Ala Val Ala Thr Val Ala Asp Pro Ala  
14 1 5 10  
15  
16 <210> 65  
17 <211> 10  
18 <212> PRT  
19 <213> Synthetic  
20  
21 <400> 65  
22  
23 Thr Val Ala Asp Pro Ala Leu Leu Ser Ile  
24 1 5 10  
25  
26 <210> 66  
27 <211> 10  
28 <212> PRT  
29 <213> Synthetic  
30  
31 <400> 66  
32  
33 Pro Ala Leu Leu Ser Ile Ser Val Ala Thr  
34 1 5 10  
35  
36 <210> 67  
37 <211> 10  
38 <212> PRT  
39 <213> Synthetic  
40  
41 <400> 67  
42  
43 Ser Ile Ser Val Ala Thr Arg Leu Ala Gln  
44 1 5 10  
45  
46 <210> 68  
47 <211> 10  
48 <212> PRT  
49 <213> Synthetic  
50  
51 <400> 68  
52  
53 Ala Thr Arg Leu Ala Gln Lys Asp Pro Ala  
54 1 5 10  
55  
56 <210> 69  
57 <211> 10

1 <212> PRT  
2 <213> Synthetic  
3  
4 <400> 69  
5  
6 Ala Gin Lys Asp Pro Ala Phe Ala Ser Ile  
7 1 5 10  
8  
9 <210> 70  
10 <211> 10  
11 <212> PRT  
12 <213> Synthetic  
13  
14 <400> 70  
15  
16 Pro Ala Phe Ala Ser Ile Arg Ser Leu Asn  
17 1 5 10  
18  
19 <210> 71  
20 <211> 10  
21 <212> PRT  
22 <213> Synthetic  
23  
24 <400> 71  
25  
26 Ser Ile Arg Ser Leu Asn Tyr Lys Lys Met  
27 1 5 10  
28  
29 <210> 72  
30 <211> 10  
31 <212> PRT  
32 <213> Synthetic  
33  
34 <400> 72  
35  
36 Leu Asn Tyr Lys Lys Met Leu Leu Leu Ser  
37 1 5 10  
38  
39 <210> 73  
40 <211> 10  
41 <212> PRT  
42 <213> Synthetic  
43  
44 <400> 73  
45  
46 Lys Met Leu Leu Leu Ser Leu Gly Thr Gly  
47 1 5 10  
48  
49 <210> 74  
50 <211> 10  
51 <212> PRT  
52 <213> Synthetic  
53  
54 <400> 74  
55  
56 Leu Ser Leu Gly Thr Gly Thr Thr Ser Glu  
57 1 5 10

1  
2 <210> 75  
3 <211> 10  
4 <212> PRT  
5 <213> Synthetic  
6  
7 <400> 75  
8  
9 Thr Gly Thr Thr Ser Glu Phe Asp Lys Thr  
10 1 5 10  
11  
12 <210> 76  
13 <211> 10  
14 <212> PRT  
15 <213> Synthetic  
16  
17 <400> 76  
18  
19 Ser Glu Phe Asp Lys Thr Tyr Thr Ala Lys  
20 1 5 10  
21  
22 <210> 77  
23 <211> 10  
24 <212> PRT  
25 <213> Synthetic  
26  
27 <400> 77  
28  
29 Lys Thr Tyr Thr Ala Lys Glu Ala Ala Thr  
30 1 5 10  
31  
32 <210> 78  
33 <211> 10  
34 <212> PRT  
35 <213> Synthetic  
36  
37 <400> 78  
38  
39 Ala Lys Glu Ala Ala Thr Trp Thr Ala Val  
40 1 5 10  
41  
42 <210> 79  
43 <211> 10  
44 <212> PRT  
45 <213> Synthetic  
46  
47 <400> 79  
48  
49 Ala Thr Trp Thr Ala Val His Trp Met Leu  
50 1 5 10  
51  
52 <210> 80  
53 <211> 10  
54 <212> PRT  
55 <213> Synthetic  
56  
57 <400> 80

1  
2 Ala Val His Trp Met Leu Val Ile Gln Lys  
3 1 5 10  
4  
5 <210> 81  
6 <211> 10  
7 <212> PRT  
8 <213> Synthetic  
9  
10 <400> 81  
11  
12 Met Leu Val Ile Gln Lys Met Thr Asp Ala  
13 1 5 10  
14  
15 <210> 82  
16 <211> 10  
17 <212> PRT  
18 <213> Synthetic  
19  
20 <400> 82  
21  
22 Gln Lys Met Thr Asp Tyr Tyr Leu Ser Thr  
23 1 5 10  
24  
25 <210> 83  
26 <211> 10  
27 <212> PRT  
28 <213> Synthetic  
29  
30 <400> 83  
31  
32 Asp Ala Ala Ser Ser Tyr Met Thr Asp Tyr  
33 1 5 10  
34  
35 <210> 84  
36 <211> 10  
37 <212> PRT  
38 <213> Synthetic  
39  
40 <400> 84  
41  
42 Ser Tyr Met Thr Asp Tyr Tyr Leu Ser Thr  
43 1 5 10  
44  
45 <210> 85  
46 <211> 10  
47 <212> PRT  
48 <213> Synthetic  
49  
50 <400> 85  
51  
52 Asp Tyr Tyr Leu Ser Thr Ala Phe Gln Ala  
53 1 5 10  
54  
55 <210> 86  
56 <211> 10  
57 <212> PRT

1 <213> Synthetic  
2  
3 <400> 86  
4  
5 Ser Thr Ala Phe Gln Ala Leu Asp Ser Lys  
6 1 5 10  
7  
8 <210> 87  
9 <211> 10  
10 <212> PRT  
11 <213> Synthetic  
12  
13 <400> 87  
14  
15 Gln Ala Leu Asp Ser Lys Asn Asn Tyr Leu  
16 1 5 10  
17  
18 <210> 88  
19 <211> 10  
20 <212> PRT  
21 <213> Synthetic  
22  
23 <400> 88  
24  
25 Ser Lys Asn Asn Tyr Leu Arg Val Gln Glu  
26 1 5 10  
27  
28 <210> 89  
29 <211> 10  
30 <212> PRT  
31 <213> Synthetic  
32  
33 <400> 89 .  
34  
35 Tyr Leu Arg Val Gln Glu Asn Ala Leu Thr  
36 1 5 10  
37  
38 <210> 90  
39 <211> 10  
40 <212> PRT  
41 <213> Synthetic  
42  
43 <400> 90  
44  
45 Gln Glu Asn Ala Leu Thr Gly Thr Thr Thr  
46 1 5 10  
47  
48 <210> 91  
49 <211> 10  
50 <212> PRT  
51 <213> Synthetic  
52  
53 <400> 91  
54  
55 Leu Thr Gly Thr Thr Thr Glu Met Asp Asp  
56 1 5 10  
57

1 <210> 92  
2 <211> 10  
3 <212> PRT  
4 <213> Synthetic  
5  
6 <400> 92  
7  
8 Thr Thr Glu Met Asp Asp Ala Ser Glu Ala  
9 1 5 10  
10  
11 <210> 93  
12 <211> 10  
13 <212> PRT  
14 <213> Synthetic  
15  
16 <400> 93  
17  
18 Asp Asp Ala Ser Glu Ala Asn Met Glu Leu  
19 1 5 10  
20  
21 <210> 94  
22 <211> 10  
23 <212> PRT  
24 <213> Synthetic  
25  
26 <400> 94  
27  
28 Glu Ala Asn Met Glu Leu Leu Val Gln Val  
29 1 5 10  
30  
31 <210> 95  
32 <211> 10  
33 <212> PRT  
34 <213> Synthetic  
35  
36 <400> 95  
37  
38 Glu Leu Leu Val Gln Val Gly Glu Asn Leu  
39 1 5 10  
40  
41 <210> 96  
42 <211> 10  
43 <212> PRT  
44 <213> Synthetic  
45  
46 <400> 96  
47  
48 Gln Val Gly Glu Asn Leu Leu Lys Lys Pro  
49 1 5 10  
50  
51 <210> 97  
52 <211> 10  
53 <212> PRT  
54 <213> Synthetic  
55  
56 <400> 97  
57

1 Asn Leu Leu Lys Lys Pro Val Ser Glu Asp  
2 1 5 10  
3  
4 <210> 98  
5 <211> 10  
6 <212> PRT  
7 <213> Synthetic  
8  
9 <400> 98  
10  
11 Lys Pro Val Ser Glu Asp Asn Pro Glu Thr  
12 1 5 10  
13  
14 <210> 99  
15 <211> 10  
16 <212> PRT  
17 <213> Synthetic  
18  
19 <400> 99  
20  
21 Glu Asp Asn Pro Glu Thr Tyr Glu Glu Ala  
22 1 5 10  
23  
24 <210> 100  
25 <211> 10  
26 <212> PRT  
27 <213> Synthetic  
28  
29 <400> 100  
30  
31 Glu Thr Tyr Glu Glu Ala Leu Lys Arg Phe  
32 1 5 10  
33  
34 <210> 101  
35 <211> 10  
36 <212> PRT  
37 <213> Synthetic  
38  
39 <400> 101  
40  
41 Glu Ala Leu Lys Arg Phe Ala Lys Leu Leu  
42 1 5 10  
43  
44 <210> 102  
45 <211> 10  
46 <212> PRT  
47 <213> Synthetic  
48  
49 <400> 102  
50  
51 Arg Phe Ala Lys Leu Leu Ser Asp Arg Lys  
52 1 5 10  
53  
54 <210> 103  
55 <211> 10  
56 <212> PRT  
57 <213> Synthetic

```

2 <400> 103
3
4 Leu Leu Ser Asp Arg Lys Lys Leu Arg Ala
5 1 . 5 10
6
7 <210> 104
8 <211> 10
9 <212> PRT
10 <213> Synthetic
11
12 <400> 104
13
14 Arg Lys Lys Leu Arg Ala Asn Lys Ala Ser
15 1 5 10
16
17 <210> 105
18 <211> 10
19 <212> PRT
20 <213> Synthetic
21
22 <400> 105
23
24 Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser
25 1 5 10
26
27 <210> 106
28 <211> 10
29 <212> PRT
30 <213> Synthetic
31
32 <400> 106
33
34 Asp Tyr Phe Asp Val Ile Ala Gly Thr Ser
35 1 5 10
36
37 <210> 107
38 <211> 10
39 <212> PRT
40 <213> Synthetic
41
42 <400> 107
43
44 Val Ile Gly Gly Thr Ser Thr Gly Gly Leu
45 1 5 10
46
47 <210> 108
48 <211> 10
49 <212> PRT
50 <213> Synthetic
51
52 <400> 108
53
54 Val Ile Ala Gly Thr Ser Thr Gly Ala Leu
55 1 5 10
56
57 <210> 109

```

1 <211> 10  
2 <212> PRT  
3 <213> Synthetic  
4  
5 <400> 109  
6  
7 Ala Phe Tyr Phe Glu His Gly Pro Gln Ile  
8 1 5 10  
9  
10 <210> 110  
11 <211> 10  
12 <212> PRT  
13 <213> Synthetic  
14  
15 <400> 110  
16  
17 Pro Ala Tyr Phe Glu His Gly Pro Gln Ile  
18 1 5 10  
19  
20 <210> 111  
21 <211> 10  
22 <212> PRT  
23 <213> Synthetic  
24  
25 <400> 111  
26  
27 Pro Phe Ala Phe Glu His Gly Pro Gln Ile  
28 1 5 10  
29  
30 <210> 112  
31 <211> 10  
32 <212> PRT  
33 <213> Synthetic  
34  
35 <400> 112  
36  
37 Pro Phe Tyr Ala Glu His Gly Pro Gln Ile  
38 1 5 10  
39  
40 <210> 113  
41 <211> 10  
42 <212> PRT  
43 <213> Synthetic  
44  
45 <400> 113  
46  
47 Pro Phe Tyr Phe Ala His Gly Pro Gln Ile  
48 1 5 10  
49  
50 <210> 114  
51 <211> 10  
52 <212> PRT  
53 <213> Synthetic  
54  
55 <400> 114  
56  
57 Pro Phe Tyr Phe Glu Ala Gly Pro Gln Ile

1 1 5 10  
2  
3 <210> 115  
4 <211> 10  
5 <212> PRT  
6 <213> Synthetic  
7  
8 <400> 115  
9  
10 Pro Phe Tyr Phe Glu His Ala Pro Gln Ile  
11 1 5 10  
12  
13 <210> 116  
14 <211> 10  
15 <212> PRT  
16 <213> Synthetic  
17  
18 <400> 116  
19  
20 Pro Phe Tyr Phe Glu His Gly Ala Gln Ile  
21 1 5 10  
22  
23 <210> 117  
24 <211> 10  
25 <212> PRT  
26 <213> Synthetic  
27  
28 <400> 117  
29  
30 Pro Phe Tyr Phe Glu His Gly Pro Ala Ile  
31 1 5 10  
32  
33 <210> 118  
34 <211> 10  
35 <212> PRT  
36 <213> Synthetic  
37  
38 <400> 118  
39  
40 Pro Phe Tyr Phe Glu His Gly Pro Gln Ala  
41 1 5 10  
42  
43 <210> 119  
44 <211> 10  
45 <212> PRT  
46 <213> Synthetic  
47  
48 <400> 119  
49  
50 Thr Phe Tyr Leu Glu Asn Gly Pro Lys Ile  
51 1 5 10  
52  
53 <210> 120  
54 <211> 10  
55 <212> PRT  
56 <213> Synthetic  
57

1 <400> 120  
2  
3 Pro Phe Phe Phe His Gly Pro Gln Ile  
4 1 5 10  
5  
6 <210> 121  
7 <211> 10  
8 <212> PRT  
9 <213> Synthetic  
10  
11 <400> 121  
12  
13 Ala Tyr Leu Met Gln Val Leu Gln Glu Lys  
14 1 5 10  
15  
16 <210> 122  
17 <211> 10  
18 <212> PRT  
19 <213> Synthetic  
20  
21 <400> 122  
22  
23 Lys Ala Leu Met Gln Val Leu Gln Glu Lys  
24 1 5 10  
25  
26 <210> 123  
27 <211> 10  
28 <212> PRT  
29 <213> Synthetic  
30  
31 <400> 123  
32  
33 Lys Tyr Ala Met Gln Val Leu Gln Glu Lys  
34 1 5 10  
35  
36 <210> 124  
37 <211> 10  
38 <212> PRT  
39 <213> Synthetic  
40  
41 <400> 124  
42  
43 Lys Tyr Leu Ala Gln Val Leu Gln Glu Lys  
44 1 5 10  
45  
46 <210> 125  
47 <211> 10  
48 <212> PRT  
49 <213> Synthetic  
50  
51 <400> 125  
52  
53 Lys Tyr Leu Met Ala Val Leu Gln Glu Lys  
54 1 5 10  
55  
56 <210> 126  
57 <211> 10

1 <212> PRT  
2 <213> Synthetic  
3  
4 <400> 126  
5  
6 Lys Tyr Leu Met Gln Ala Leu Gln Glu Lys  
7 1 5 10  
8  
9 <210> 127  
10 <211> 10  
11 <212> PRT  
12 <213> Synthetic  
13  
14 <400> 127  
15  
16 Lys Tyr Leu Met Gln Val Ala Gln Glu Lys  
17 1 5 10  
18  
19 <210> 128  
20 <211> 10  
21 <212> PRT  
22 <213> Synthetic  
23  
24 <400> 128  
25  
26 Lys Tyr Leu Met Gln Val Leu Ala Glu Lys  
27 1 5 10  
28  
29 <210> 129  
30 <211> 10  
31 <212> PRT  
32 <213> Synthetic  
33  
34 <400> 129  
35  
36 Lys Tyr Leu Met Gln Val Leu Gln Ala Lys  
37 1 5 10  
38  
39 <210> 130  
40 <211> 10  
41 <212> PRT  
42 <213> Synthetic  
43  
44 <400> 130  
45  
46 Lys Tyr Leu Met Gln Val Leu Gln Glu Ala  
47 1 5 10  
48  
49 <210> 131  
50 <211> 10  
51 <212> PRT  
52 <213> Synthetic  
53  
54 <400> 131  
55  
56 Val Phe Leu His Asp Lys Ile Lys Ser Leu  
57 1 5 10

1  
2 <210> 132  
3 <211> 10  
4 <212> PRT  
5 <213> Synthetic  
6  
7 <400> 132  
8  
9 Ala Tyr Ser Thr Ala Ala Ala Pro Thr Tyr  
10 1 5 10  
11  
12 <210> 133  
13 <211> 10  
14 <212> PRT  
15 <213> Synthetic  
16  
17 <400> 133  
18  
19 Ser Ala Ser Thr Ala Ala Ala Pro Thr Tyr  
20 1 5 10  
21  
22 <210> 134  
23 <211> 10  
24 <212> PRT  
25 <213> Synthetic  
26  
27 <400> 134  
28  
29 Ser Tyr Ala Thr Ala Ala Ala Pro Thr Tyr  
30 1 5 10  
31  
32 <210> 135  
33 <211> 10  
34 <212> PRT  
35 <213> Synthetic  
36  
37 <400> 135  
38  
39 Ser Tyr Ser Ala Ala Ala Ala Pro Thr Tyr  
40 1 5 10  
41  
42 <210> 136  
43 <211> 10  
44 <212> PRT  
45 <213> Synthetic  
46  
47 <400> 136  
48  
49 Ser Tyr Ser Thr Ala Ala Ala Ala Thr Tyr  
50 1 5 10  
51  
52 <210> 137  
53 <211> 10  
54 <212> PRT  
55 <213> Synthetic  
56  
57 <400> 137

1  
2   Ser Tyr Ser Thr Ala Ala Ala Pro Ala Tyr  
3   1                   5                   10  
4  
5   <210> 138  
6   <211> 10  
7   <212> PRT  
8   <213> Synthetic  
9  
10 <400> 138  
11  
12   Ser Tyr Ser Thr Ala Ala Ala Pro Thr Ala  
13   1                   5                   10  
14  
15   <210> 139  
16   <211> 10  
17   <212> PRT  
18   <213> Synthetic  
19  
20 <400> 139  
21  
22   Cys Ile Ser Thr Ser Ala Ala Pro Thr Tyr  
23   1                   5                   10  
24  
25   <210> 140  
26   <211> 10  
27   <212> PRT  
28   <213> Synthetic  
29  
30 <400> 140  
31  
32   Ser Tyr Ser Thr Ala Ala Ala Pro Ala Phe  
33   1                   5                   10  
34  
35   <210> 141  
36   <211> 10  
37   <212> PRT  
38   <213> Synthetic  
39  
40 <400> 141  
41  
42   Ala Phe Ala Ala Ala Ala Ala Pro Thr Tyr  
43   1                   5                   10  
44  
45   <210> 142  
46   <211> 10  
47   <212> PRT  
48   <213> Synthetic  
49  
50 <400> 142  
51  
52   Ser Tyr Ser Thr Ala Ala Ala Pro Thr Phe  
53   1                   5                   10  
54  
55   <210> 143  
56   <211> 10  
57   <212> PRT

1 <213> Synthetic  
2  
3 <400> 143  
4  
5 Ser Thr Ser Ala Ala Pro Thr Tyr Phe Pro  
6 1 5 10  
7  
8 <210> 144  
9 <211> 10  
10 <212> PRT  
11 <213> Synthetic  
12  
13 <400> 144  
14  
15 Ser Thr Ser Ala Ala Pro Thr Phe Phe Pro  
16 1 5 10  
17  
18 <210> 145  
19 <211> 10  
20 <212> PRT  
21 <213> Synthetic  
22  
23 <400> 145  
24  
25 Ser Thr Ser Ala Ala Pro Thr Ala Phe Pro  
26 1 5 10  
27  
28 <210> 146  
29 <211> 10  
30 <212> PRT  
31 <213> Synthetic  
32  
33 <400> 146  
34  
35 Ser Thr Ala Ala Ala Pro Thr Phe Phe Pro  
36 1 5 10  
37  
38 <210> 147  
39 <211> 10  
40 <212> PRT  
41 <213> Synthetic  
42  
43 <400> 147  
44  
45 Ala Ala Ala Ala Thr Tyr Phe Pro Pro His  
46 1 5 10  
47  
48 <210> 148  
49 <211> 10  
50 <212> PRT  
51 <213> Synthetic  
52  
53 <400> 148  
54  
55 Ala Ala Ala Pro Ala Tyr Phe Pro Pro His  
56 1 5 10  
57

1 <210> 149  
2 <211> 10  
3 <212> PRT  
4 <213> Synthetic  
5  
6 <400> 149  
7  
8 Ala Ala Ala Pro Thr Ala Phe Pro Pro His  
9 1 5 10  
10  
11 <210> 150  
12 <211> 10  
13 <212> PRT  
14 <213> Synthetic  
15  
16 <400> 150  
17  
18 Ala Ala Ala Pro Thr Tyr Ala Pro Pro His  
19 1 5 10  
20  
21 <210> 151  
22 <211> 10  
23 <212> PRT  
24 <213> Synthetic  
25  
26 <400> 151  
27  
28 Ala Ala Ala Pro Thr Tyr Phe Ala Pro His  
29 1 5 10  
30  
31 <210> 152  
32 <211> 10  
33 <212> PRT  
34 <213> Synthetic  
35  
36 <400> 152  
37  
38 Ala Ala Ala Pro Thr Tyr Phe Pro Ala His  
39 1 5 10  
40  
41 <210> 153  
42 <211> 10  
43 <212> PRT  
44 <213> Synthetic  
45  
46 <400> 153  
47  
48 Ala Ala Ala Pro Thr Tyr Phe Pro Pro Ala  
49 1 5 10  
50  
51 <210> 154  
52 <211> 10  
53 <212> PRT  
54 <213> Synthetic  
55  
56 <400> 154  
57

1 Ser Ala Ala Pro Thr Tyr Phe Pro Ala His  
2 1 5 10  
3  
4 <210> 155  
5 <211> 10  
6 <212> PRT  
7 <213> Synthetic  
8  
9 <400> 155  
10  
11 Ala Ala Ala Pro Ala Phe Phe Pro Pro His  
12 1 5 10  
13  
14 <210> 156  
15 <211> 10  
16 <212> PRT  
17 <213> Synthetic  
18  
19 <400> 156  
20  
21 Ala Ala Ala Pro Pro Phe Phe Pro Pro His  
22 1 5 10  
23  
24 <210> 157  
25 <211> 10  
26 <212> PRT  
27 <213> Synthetic  
28  
29 <400> 157  
30  
31 Ala Ala Ala Pro Thr Phe Phe Pro Pro His  
32 1 5 10  
33  
34 <210> 158  
35 <211> 10  
36 <212> PRT  
37 <213> Synthetic  
38  
39 <400> 158  
40  
41 Ser Ile Ser Val Ala Thr Arg Leu Ala Gln  
42 1 5 10  
43  
44 <210> 159  
45 <211> 10  
46 <212> PRT  
47 <213> Synthetic  
48  
49 <400> 159  
50  
51 Ala Met Ser Met Leu Thr Lys Glu Val His  
52 1 5 10  
53  
54 <210> 160  
55 <211> 10  
56 <212> PRT  
57 <213> Synthetic

1  
2 <400> 160  
3  
4 Pro Ala Phe Ala Ser Ile Arg Ser Leu Asn  
5 1 5 10  
6  
7 <210> 161  
8 <211> 10  
9 <212> PRT  
10 <213> Synthetic  
11  
12 <400> 161  
13  
14 Pro Asn Phe Asn Ala Gly Ser Pro Thr Glu  
15 1 5 10  
16  
17 <210> 162  
18 <211> 10  
19 <212> PRT  
20 <213> Synthetic  
21  
22 <400> 162  
23  
24 Lys Met Leu Leu Leu Ser Leu Gly Thr Gly  
25 1 5 10  
26  
27 <210> 163  
28 <211> 10  
29 <212> PRT  
30 <213> Synthetic  
31  
32 <400> 163  
33  
34 Asn Tyr Leu Ile Ile Ser Val Gly Thr Gly  
35 1 5 10  
36  
37 <210> 164  
38 <211> 10  
39 <212> PRT  
40 <213> Synthetic  
41  
42 <400> 164  
43  
44 Lys Met Leu Leu Leu Ser Leu Gly Ala Gly  
45 1 5 10  
46  
47 <210> 165  
48 <211> 10  
49 <212> PRT  
50 <213> Synthetic  
51  
52 <400> 165  
53  
54 Ala Glu Phe Asp Lys Thr Tyr Thr Ala Lys  
55 1 5 10  
56  
57 <210> 166

1 <211> 10  
2 <212> PRT  
3 <213> Synthetic  
4  
5 <400> 166  
6  
7 Ser Ala Phe Asp Lys Thr Tyr Thr Ala Lys  
8 1 5 10  
9  
10 <210> 167  
11 <211> 10  
12 <212> PRT  
13 <213> Synthetic  
14  
15 <400> 167  
16  
17 Ser Glu Ala Asp Lys Thr Tyr Thr Ala Lys  
18 1 5 10  
19  
20 <210> 168  
21 <211> 10  
22 <212> PRT  
23 <213> Synthetic  
24  
25 <400> 168  
26  
27 Ser Glu Phe Ala Lys Thr Tyr Thr Ala Lys  
28 1 5 10  
29  
30 <210> 169  
31 <211> 10  
32 <212> PRT  
33 <213> Synthetic  
34  
35 <400> 169  
36  
37 Ser Glu Phe Asp Ala Thr Tyr Thr Ala Lys  
38 1 5 10  
39  
40 <210> 170  
41 <211> 10  
42 <212> PRT  
43 <213> Synthetic  
44  
45 <400> 170  
46  
47 Ser Glu Phe Asp Lys Ala Tyr Thr Ala Lys  
48 1 5 10  
49  
50 <210> 171  
51 <211> 10  
52 <212> PRT  
53 <213> Synthetic  
54  
55 <400> 171  
56  
57 Ser Glu Phe Asp Lys Thr Ala Thr Ala Lys

1 1 5 10  
2  
3 <210> 172  
4 <211> 10  
5 <212> PRT  
6 <213> Synthetic  
7  
8 <400> 172  
9  
10 Ser Glu Phe Asp Lys Thr Tyr Ala Ala Lys  
11 1 5 10  
12  
13 <210> 173  
14 <211> 10  
15 <212> PRT  
16 <213> Synthetic  
17  
18 <400> 173  
19  
20 Ser Glu Phe Asp Lys Thr Tyr Thr Ala Ala  
21 1 5 10  
22  
23 <210> 174  
24 <211> 10  
25 <212> PRT  
26 <213> Synthetic  
27  
28 <400> 174  
29  
30 Lys Gln Ala Glu Lys Tyr Thr Ala Glu Gln  
31 1 5 10  
32  
33 <210> 175  
34 <211> 10  
35 <212> PRT  
36 <213> Synthetic  
37  
38 <400> 175  
39  
40 Ser Glu Phe Asp Ala Ala Phe Ala Ala Ala  
41 1 5 10  
42  
43 <210> 176  
44 <211> 10  
45 <212> PRT  
46 <213> Synthetic  
47  
48 <400> 176  
49  
50 Ser Glu Phe Asp Lys Thr Phe Thr Ala Lys  
51 1 5 10  
52  
53 <210> 177  
54 <211> 10  
55 <212> PRT  
56 <213> Synthetic  
57

1 <400> 177  
2  
3 Ala Glu Lys Tyr Thr Ala Glu Gln Cys Ala  
4 1 5 10  
5  
6 <210> 178  
7 <211> 10  
8 <212> PRT  
9 <213> Synthetic  
10  
11 <400> 178  
12  
13 Ala Thr Tyr Thr Ala Lys Glu Ala Ala Thr  
14 1 5 10  
15  
16 <210> 179  
17 <211> 10  
18 <212> PRT  
19 <213> Synthetic  
20  
21 <400> 179  
22  
23 Lys Ala Tyr Thr Ala Lys Glu Ala Ala Thr  
24 1 5 10  
25  
26 <210> 180  
27 <211> 10  
28 <212> PRT  
29 <213> Synthetic  
30  
31 <400> 180  
32  
33 Lys Thr Ala Thr Ala Lys Glu Ala Ala Thr  
34 1 5 10  
35  
36 <210> 181  
37 <211> 10  
38 <212> PRT  
39 <213> Synthetic  
40  
41 <400> 181  
42  
43 Lys Thr Tyr Ala Ala Lys Glu Ala Ala Thr  
44 1 5 10  
45  
46 <210> 182  
47 <211> 10  
48 <212> PRT  
49 <213> Synthetic  
50  
51 <400> 182  
52  
53 Lys Thr Tyr Thr Ala Ala Glu Ala Ala Thr  
54 1 5 10  
55  
56 <210> 183  
57 <211> 10

1 <212> PRT  
2 <213> Synthetic  
3  
4 <400> 183  
5  
6 Lys Thr Tyr Thr Ala Lys Ala Ala Ala Thr  
7 1 5 10  
8  
9 <210> 184  
10 <211> 10  
11 <212> PRT  
12 <213> Synthetic  
13  
14 <400> 184  
15  
16 Lys Thr Tyr Thr Ala Lys Glu Ala Ala Ala  
17 1 5 10  
18  
19 <210> 185  
20 <211> 10  
21 <212> PRT  
22 <213> Synthetic  
23  
24 <400> 185  
25  
26 Glu Lys Tyr Thr Ala Glu Gln Cys Ala Lys  
27 1 5 10  
28  
29 <210> 186  
30 <211> 10  
31 <212> PRT  
32 <213> Synthetic  
33  
34 <400> 186  
35  
36 Ala Ala Phe Ala Ala Ala Glu Ala Ala Thr  
37 1 5 10  
38  
39 <210> 187  
40 <211> 10  
41 <212> PRT  
42 <213> Synthetic  
43  
44 <400> 187  
45  
46 Lys Thr Phe Thr Ala Lys Glu Ala Ala Thr  
47 1 5 10  
48  
49 <210> 188  
50 <211> 10  
51 <212> PRT  
52 <213> Synthetic  
53  
54 <400> 188  
55  
56 Gln Ala Leu His Cys Glu Lys Lys Tyr Leu  
57 1 5 10

1  
2 <210> 189  
3 <211> 10  
4 <212> PRT  
5 <213> Synthetic  
6  
7 <400> 189  
8  
9 Gln Ala Leu Asp Ser Lys Ala Ala Tyr Leu  
10 1 5 10  
11  
12 <210> 190  
13 <211> 10  
14 <212> PRT  
15 <213> Synthetic  
16  
17 <400> 190  
18  
19 Gln Ala Leu Asp Ser Lys Asn Asn Phe Leu  
20 1 5 10  
21  
22 <210> 191  
23 <211> 10  
24 <212> PRT  
25 <213> Synthetic  
26  
27 <400> 191  
28  
29 Gln Ala Leu His Cys Glu Asn Asn Phe Leu  
30 1 5 10  
31  
32 <210> 192  
33 <211> 10  
34 <212> PRT  
35 <213> Synthetic  
36  
37 <400> 192  
38  
39 Cys Glu Lys Lys Tyr Leu Arg Ile Gln Asp  
40 1 5 10  
41  
42 <210> 193  
43 <211> 10  
44 <212> PRT  
45 <213> Synthetic  
46  
47 <400> 193  
48  
49 Ser Lys Asn Asn Phe Leu Arg Val Gln Glu  
50 1 5 10  
51  
52 <210> 194  
53 <211> 10  
54 <212> PRT  
55 <213> Synthetic  
56  
57 <400> 194

1  
2   Ser Glu Asn Asn Tyr Leu Arg Val Gln Glu  
3   1    10  
4  
5   <210> 195  
6   <211> 10  
7   <212> PRT  
8   <213> Synthetic  
9  
10   <400> 195  
11  
12   Ala Leu Arg Val Gln Glu Asn Ala Leu Thr  
13   1    10  
14  
15   <210> 196  
16   <211> 10  
17   <212> PRT  
18   <213> Synthetic  
19  
20   <400> 196  
21  
22   Tyr Ala Arg Val Gln Glu Asn Ala Leu Thr  
23   1    10  
24  
25   <210> 197  
26   <211> 10  
27   <212> PRT  
28   <213> Synthetic  
29  
30   <400> 197  
31  
32   Tyr Leu Ala Val Gln Glu Asn Ala Leu Thr  
33   1    10  
34  
35   <210> 198  
36   <211> 10  
37   <212> PRT  
38   <213> Synthetic  
39  
40   <400> 198  
41  
42   Tyr Leu Arg Ala Gln Glu Asn Ala Leu Thr  
43   1    10  
44  
45   <210> 199  
46   <211> 10  
47   <212> PRT  
48   <213> Synthetic  
49  
50   <400> 199  
51  
52   Tyr Leu Arg Val Ala Glu Asn Ala Leu Thr  
53   1    10  
54  
55   <210> 200  
56   <211> 10  
57   <212> PRT

1 <213> Synthetic  
2  
3 <400> 200  
4  
5 Tyr Leu Arg Val Gln Ala Asn Ala Leu Thr  
6 1 5 10  
7  
8 <210> 201  
9 <211> 10  
10 <212> PRT  
11 <213> Synthetic  
12  
13 <400> 201  
14  
15 Tyr Leu Arg Val Gln Glu Ala Ala Leu Thr  
16 1 5 10  
17  
18 <210> 202  
19 <211> 10  
20 <212> PRT  
21 <213> Synthetic  
22  
23 <400> 202  
24  
25 Tyr Leu Arg Val Gln Glu Asn Ala Ala Thr  
26 1 5 10  
27  
28 <210> 203  
29 <211> 10  
30 <212> PRT  
31 <213> Synthetic  
32  
33 <400> 203  
34  
35 Tyr Leu Arg Val Gln Glu Asn Ala Leu Ala  
36 1 5 10  
37  
38 <210> 204  
39 <211> 10  
40 <212> PRT  
41 <213> Synthetic  
42  
43 <400> 204  
44  
45 Tyr Leu Arg Ile Gln Asp Asp Thr Leu Thr  
46 1 5 10  
47  
48 <210> 205  
49 <211> 10  
50 <212> PRT  
51 <213> Synthetic  
52  
53 <400> 205  
54  
55 Tyr Leu Thr Val Ala Ala Ala Ala Leu Thr  
56 1 5 10  
57

1 <210> 206  
2 <211> 10  
3 <212> PRT  
4 <213> Synthetic  
5  
6 <400> 206  
7  
8 Phe Leu Arg Val Gln Glu Asn Ala Leu Thr  
9 1 5 10  
10  
11 <210> 207  
12 <211> 10  
13 <212> PRT  
14 <213> Synthetic  
15  
16 <400> 207  
17  
18 Asn Asn Tyr Leu Arg Val Gln Glu Asn Ala  
19 1 5 10  
20  
21 <210> 208  
22 <211> 10  
23 <212> PRT  
24 <213> Synthetic  
25  
26 <400> 208  
27  
28 Lys Lys Tyr Leu Arg Ile Gln Asp Asp Thr  
29 1 5 10  
30  
31 <210> 209  
32 <211> 10  
33 <212> PRT  
34 <213> Synthetic  
35  
36 <400> 209  
37  
38 Asn Asn Phe Leu Arg Val Gln Glu Asn Ala  
39 1 5 10  
40  
41 <210> 210  
42 <211> 10  
43 <212> PRT  
44 <213> Synthetic  
45  
46 <400> 210  
47  
48 Asn Ala Tyr Leu Arg Val Gln Glu Asn Ala  
49 1 5 10  
50  
51 <210> 211  
52 <211> 10  
53 <212> PRT  
54 <213> Synthetic  
55  
56 <400> 211  
57

1 Ala Thr Tyr Glu Glu Ala Lys Leu Arg Phe  
2 1 5 10  
3  
4 <210> 212  
5 <211> 10  
6 <212> PRT  
7 <213> Synthetic  
8  
9 <400> 212  
10  
11 Glu Ala Tyr Glu Glu Ala Leu Lys Arg Phe  
12 1 5 10  
13  
14 <210> 213  
15 <211> 10  
16 <212> PRT  
17 <213> Synthetic  
18  
19 <400> 213  
20  
21 Glu Thr Ala Glu Glu Ala Leu Lys Arg Phe  
22 1 5 10  
23  
24 <210> 214  
25 <211> 10  
26 <212> PRT  
27 <213> Synthetic  
28  
29 <400> 214  
30  
31 Glu Thr Tyr Ala Glu Ala Leu Lys Arg Phe  
32 1 5 10  
33  
34 <210> 215  
35 <211> 10  
36 <212> PRT  
37 <213> Synthetic  
38  
39 <400> 215  
40  
41 Glu Thr Tyr Glu Ala Ala Leu Lys Arg Phe  
42 1 5 10  
43  
44 <210> 216  
45 <211> 10  
46 <212> PRT  
47 <213> Synthetic  
48  
49 <400> 216  
50  
51 Glu Thr Tyr Glu Glu Ala Ala Lys Arg Phe  
52 1 5 10  
53  
54 <210> 217  
55 <211> 10  
56 <212> PRT  
57 <213> Synthetic

1  
2 <400> 217  
3  
4 Glu Thr Tyr Glu Glu Ala Leu Ala Arg Phe  
5 1 5 10  
6  
7 <210> 218  
8 <211> 10  
9 <212> PRT  
10 <213> Synthetic  
11  
12 <400> 218  
13  
14 Glu Thr Tyr Glu Glu Ala Leu Lys Ala Phe  
15 1 5 10  
16  
17 <210> 219  
18 <211> 10  
19 <212> PRT  
20 <213> Synthetic  
21  
22 <400> 219  
23  
24 Glu Thr Tyr Glu Glu Ala Leu Lys Arg Ala  
25 1 5 10  
26  
27 <210> 220  
28 <211> 10  
29 <212> PRT  
30 <213> Synthetic  
31  
32 <400> 220  
33  
34 Gly Thr Asn Ala Gln Ser Leu Ala Asp Phe  
35 1 5 10  
36  
37 <210> 221  
38 <211> 10  
39 <212> PRT  
40 <213> Synthetic  
41  
42 <400> 221  
43  
44 Glu Thr Tyr Glu Ala Ala Leu Ala Ala Phe  
45 1 5 10  
46  
47 <210> 222  
48 <211> 10  
49 <212> PRT  
50 <213> Synthetic  
51  
52 <400> 222  
53  
54 Glu Thr Phe Glu Glu Ala Leu Lys Arg Phe  
55 1 5 10  
56  
57 <210> 223

1 <211> 10  
2 <212> PRT  
3 <213> Synthetic  
4  
5 <400> 223  
6  
7 Tyr Glu Glu Ala Leu Lys Thr Phe Ala Lys  
8 1 5 10  
9  
10 <210> 224  
11 <211> 10  
12 <212> PRT  
13 <213> Synthetic  
14  
15 <400> 224  
16  
17 Phe Glu Glu Ala Leu Lys Arg Phe Ala Lys  
18 1 5 10  
19  
20 <210> 225  
21 <211> 10  
22 <212> PRT  
23 <213> Synthetic  
24  
25 <400> 225  
26  
27 Ala Ala Leu Lys Arg Phe Ala Lys Leu Leu  
28 1 5 10  
29  
30 <210> 226  
31 <211> 10  
32 <212> PRT  
33 <213> Synthetic  
34  
35 <400> 226  
36  
37 Glu Ala Ala Lys Arg Phe Ala Lys Leu Leu  
38 1 5 10  
39  
40 <210> 227  
41 <211> 10  
42 <212> PRT  
43 <213> Synthetic  
44  
45 <400> 227  
46  
47 Glu Ala Leu Ala Arg Phe Ala Lys Leu Leu  
48 1 5 10  
49  
50 <210> 228  
51 <211> 10  
52 <212> PRT  
53 <213> Synthetic  
54  
55 <400> 228  
56  
57 Glu Ala Leu Lys Ala Phe Ala Lys Leu Leu

1 1 5 10  
2  
3 <210> 229  
4 <211> 10  
5 <212> PRT  
6 <213> Synthetic  
7  
8 <400> 229  
9  
10 Glu Ala Leu Lys Arg Ala Ala Lys Leu Leu  
11 1 5 10  
12  
13 <210> 230  
14 <211> 10  
15 <212> PRT  
16 <213> Synthetic  
17  
18 <400> 230  
19  
20 Glu Ala Leu Lys Arg Phe Ala Ala Leu Leu  
21 1 5 10  
22  
23 <210> 231  
24 <211> 10  
25 <212> PRT  
26 <213> Synthetic  
27  
28 <400> 231  
29  
30 Glu Ala Leu Lys Arg Phe Ala Lys Ala Leu  
31 1 5 10  
32  
33 <210> 232  
34 <211> 10  
35 <212> PRT  
36 <213> Synthetic  
37  
38 <400> 232  
39  
40 Glu Ala Leu Lys Arg Phe Ala Lys Leu Ala  
41 1 5 10  
42  
43 <210> 233  
44 <211> 10  
45 <212> PRT  
46 <213> Synthetic  
47  
48 <400> 233  
49  
50 Gln Ser Leu Ala Asp Phe Ala Lys Gln Leu  
51 1 5 10  
52  
53 <210> 234  
54 <211> 10  
55 <212> PRT  
56 <213> Synthetic  
57

1 <400> 234  
2  
3 Ala Ala Leu Ala Ala Phe Ala Lys Leu Leu  
4 1 5 10  
5  
6 <210> 235  
7 <211> 10  
8 <212> PRT  
9 <213> Synthetic  
10  
11 <400> 235  
12  
13 Leu Ala Asp Phe Ala Lys Gln Leu Ser Asp  
14 1 5 10  
15  
16 <210> 236  
17 <211> 10  
18 <212> PRT  
19 <213> Synthetic  
20  
21 <400> 236  
22  
23 Asp Phe Ala Lys Gln Leu Ser Asp Glu Arg  
24 1 5 10  
25  
26 <210> 237  
27 <211> 10  
28 <212> PRT  
29 <213> Synthetic  
30  
31 <400> 237  
32  
33 Ala Phe Ala Ala Leu Leu Ser Asp Arg Lys  
34 1 5 10  
35  
36 <210> 238  
37 <211> 10  
38 <212> PRT  
39 <213> Synthetic  
40  
41 <400> 238  
42  
43 Ser Thr Ala Ala Ala Pro Thr Tyr Phe Pro  
44 1 5 10  
45  
46 <210> 239  
47 <211> 10  
48 <212> PRT  
49 <213> Synthetic  
50  
51 <400> 239  
52  
53 Leu Lys Arg Phe Ala Lys Leu Leu Ser Asp  
54 1 5 10  
55  
56 <210> 240  
57 <211> 10

```

1 <212> PRT
2 <213> Synthetic
3
4 <400> 240
5
6 Gln Ala Leu Asp Ser Glu Asn Asn Phe Leu
7 1 5 10
8
9 <210> 241
10 <211> 10
11 <212> PRT
12 <213> Synthetic
13
14 <400> 241
15
16 Ser Asp Leu Ala Asp Phe Ala Lys Gln Leu
17 1 5 10
18
19 <210> 242
20 <211> 55
21 <212> DNA
22 <213> Synthetic
23
24 <400> 242
25 ggagctcgag aaaagagagg ctgaagcttc attgaattac aaaaaatgc tgttg 55
26
27
28 <210> 243
29 <211> 42
30 <212> DNA
31 <213> Synthetic
32
33 <400> 243
34 tcccaactgt cctggccat aagaagctt gtttgcgg ag 42
35
36
37 <210> 244
38 <211> 36
39 <212> DNA
40 <213> Synthetic
41
42 <400> 244
43 gtttctatg gaccaggaca gttggagaa atggta 36
44
45 <210> 245
46 <211> 39
47 <212> DNA
48 <213> Synthetic
49
50 <400> 245
51 ggcttagagg aattctcatt acctaattga agcaaatgc 39
52
53 <210> 246
54 <211> 1128
55 <212> DNA
56 <213> Synthetic
57

```

1 <400> 246  
 2 tcgagaaaag agaggctgaa gcttcattga attacaaaaa aatgctgttgc tcttcattag 60  
 3  
 4 gcactggcac tacttcagag tttgataaaa catatacagc aaaagaggca gctacctgga 120  
 5  
 6 ctgtctaca ttggatgttata gttatacaga aaatgactga tgcagcaagt tcttacatga 180  
 7  
 8 ctgttattttact gctttcaag ctcttgattc aaaaacaat tacctcaggg 240  
 9  
 10 ttcaagaaaaa tgcattaaca ggcacaacta ctgaaatggaa tgatgcttct gaggctaata 300  
 11  
 12 tggatttattt agtacaagtt ggtgaaaact tattgaaagaa accagttcc gaagacaatc 360  
 13  
 14 ctgaaaccta tgaggaagct ctaaagaggt ttgcaaaattt gctctctgtt aggaagaaac 420  
 15  
 16 tccgagcaaa caaaagcttct tatggaccag gacagttggg agaaatggt actgttctta 480  
 17  
 18 gtattgttggg aggttggattt aggggatca ttccggctac catttcgaa tttttttaag 540  
 19  
 20 gacaactca gggaaatggac aataatgcag atgcaagact tgcagattac ttttgcattaa 600  
 21  
 22 ttggaggaac aagtacagga ggtttatttga ctgctatgtt aagtactcca aatgaaaaca 660  
 23  
 24 atcgaccctt tgctgctgcc aaagaaattt taccccttttta cttcgaacat ggccttcaga 720  
 25  
 26 tttttatcc tagtggtaaa attttaggcc caaaatatga tggaaaatataat cttatgcag 780  
 27  
 28 ttcttcaaga aaaaacttggaa gaaactctgtg tgcatcaagc tttgacagaa gttgtcatct 840  
 29  
 30 caagctttga cataaaaaaca aataagccag taatattcataatc taagtcaaat ttagcaact 900  
 31  
 32 ctccagaattt ggatgctaaatgatgaca taagttatttcc cacacggca gctccaaat 960  
 33  
 34 atttcccttgc cattactttt gttactaata ctgatataatgg agatgaaatat gagttcaatc 1020  
 35  
 36 ttgttgcatttgc tgctgttgc actgttgcgtt atccggcggtt attatccattt akggttgc 1080  
 37  
 38  
 39 cgagacttgc acaaaaaggat ccagcatttg cttaatttgc gtaatgag 1128  
 40  
 41 <210> 247  
 42 <211> 366  
 43 <212> PRT  
 44 <213> Synthetic  
 45  
 46 <400> 247  
 47  
 48 Ser Leu Asn Tyr Lys Lys Met Leu Leu Leu Ser Leu Gly Thr Gly Thr  
 49 1 5 10 15  
 50  
 51 Thr Ser Glu Phe Asp Lys Thr Tyr Thr Ala Lys Glu Ala Ala Thr Trp  
 52 20 25 30  
 53  
 54 Thr Ala Val His Trp Met Leu Val Ile Gln Lys Met Thr Asp Ala Ala  
 55 35 40 45  
 56  
 57 Ser Ser Tyr Met Thr Asp Tyr Tyr Leu Ser Thr Ala Phe Gln Ala Leu

1	50	55	60	
2	Asp Ser Lys Asn Asn Tyr Leu Arg Val Gln Glu Asn Ala Leu Thr Gly			
3	65	70	75	80
4				
5	Thr Thr Thr Glu Met Asp Asp Ala Ser Glu Ala Asn Met Glu Leu Leu			
6	85	90		95
7				
8	Val Gin Val Gly Glu Asn Leu Leu Lys Lys Pro Val Ser Glu Asp Asn			
9	100	105		110
10				
11	Pro Glu Thr Tyr Glu Glu Ala Leu Lys Arg Phe Ala Lys Leu Leu Ser			
12	115	120		125
13				
14	Asp Arg Lys Lys Leu Arg Ala Asn Lys Ala Ser Tyr Gly Pro Gly Gln			
15	130	135		140
16				
17	Leu Gly Glu Met Val Thr Val Leu Ser Ile Asp Gly Gly Gly Ile Arg			
18	145	150		160
19				
20	Gly Ile Ile Pro Ala Thr Ile Leu Glu Phe Leu Glu Gly Gln Leu Gln			
21	165	170		175
22				
23	Glu Met Asp Asn Asn Ala Asp Ala Arg Leu Ala Asp Tyr Phe Asp Val			
24	180	185		190
25				
26	Ile Gly Gly Thr Ser Thr Gly Gly Leu Leu Thr Ala Met Ile Ser Thr			
27	195	200		205
28				
29	Pro Asn Glu Asn Asn Arg Pro Phe Ala Ala Ala Lys Glu Ile Val Pro			
30	210	215		220
31				
32	Phe Tyr Phe Glu His Gly Pro Gln Ile Phe Asn Pro Ser Gly Gln Ile			
33	225	230		235
34				240
35	Leu Gly Pro Lys Tyr Asp Gly Lys Tyr Leu Met Gln Val Leu Gln Glu			
36	245	250		255
37				
38	Lys Leu Gly Glu Thr Arg Val His Gln Ala Leu Thr Glu Val Val Ile			
39	260	265		270
40				
41	Ser Ser Phe Asp Ile Lys Thr Asn Lys Pro Val Ile Phe Thr Lys Ser			
42	275	280		285
43				
44	Asn Leu Ala Asn Ser Pro Glu Leu Asp Ala Lys Met Tyr Asp Ile Ser			
45	290	295		300
46				
47	Tyr Ser Thr Ala Ala Ala Pro Thr Tyr Phe Pro Pro His Tyr Phe Val			
48	305	310		320
49				
50	Thr Asn Thr Ser Asn Gly Asp Glu Tyr Glu Phe Asn Leu Val Asp Gly			
51	325	330		335
52				
53	Ala Val Ala Thr Val Ala Asp Pro Ala Leu Leu Ser Ile Ser Val Ala			
54	340	345		350
55				
56	Thr Arg Leu Ala Gln Lys Asp Pro Ala Phe Ala Ser Ile Arg			
57				

1	355	360	365
2	<210> 248		
3	<211> 55		
4	<212> DNA		
5	<213> Synthetic		
6			
7	<400> 248		
8	ggagctcgag aaaagagagg ctgaagctaa tactagtaat ggagatgaat atgag		55
9			
10	<210> 249		
11	<211> 39		
12	<212> DNA		
13	<213> Synthetic		
14			
15	<400> 249		
16	ggcttagagg aattctcatt aagaataaaaa gtaatgcgg		39
17			
18	<210> 250		
19	<211> 1128		
20	<212> DNA		
21	<213> Synthetic		
22			
23	<400> 250		
24	tcgagaaaaag agaggctgaa gctataacta gtaatggaga tgaatatgag ttcaatcttg		60
25	tttatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		120
26	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		180
27	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		240
28	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		300
29	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		360
30	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		420
31	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		480
32	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		540
33	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		600
34	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		660
35	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		720
36	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		780
37	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		840
38	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		900
39	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		960
40	ttgtatggtgc ttgttctact ttgtctgatc cggcgattt atccatttgc gtttcaacgc		1020

1 caaatttagc aaactcttcca gaattggatg ctaagatgt a tgacataagt tattccacag 1080  
2  
3 cagcagctcc aacatatttt cctccgcatt acttggttac ttaatgag 1128  
4  
5 <210> 251  
6 <211> 366  
7 <212> PRT  
8 <213> Synthetic  
9  
10 <400> 251  
11  
12 Asn Thr Ser Asn Gly Asp Glu Tyr Glu Phe Asn Leu Val Asp Gly Ala  
13 1 5 10 15  
14  
15 Val Ala Thr Val Ala Asp Pro Ala Leu Leu Ser Ile Ser Val Ala Thr  
16 20 25 30  
17  
18 Arg Leu Ala Gln Lys Asp Pro Ala Phe Ala Ser Ile Arg Ser Leu Asn  
19 35 40 45  
20  
21 Tyr Lys Lys Met Leu Leu Leu Ser Leu Gly Thr Gly Thr Thr Ser Glu  
22 50 55 60  
23  
24 Phe Asp Lys Thr Tyr Thr Ala Lys Glu Ala Ala Thr Trp Thr Ala Val  
25 65 70 75 80  
26  
27 His Trp Met Leu Val Ile Gln Lys Met Thr Asp Ala Ala Ser Ser Tyr  
28 85 90 95  
29  
30 Met Thr Asp Tyr Tyr Leu Ser Thr Ala Phe Gln Ala Leu Asp Ser Lys  
31 100 105 110  
32  
33 Asn Asn Tyr Leu Arg Val Gln Glu Asn Ala Leu Thr Gly Thr Thr Thr  
34 115 120 125  
35  
36 Glu Met Asp Asp Ala Ser Glu Ala Asn Met Glu Leu Leu Val Gln Val  
37 130 135 140  
38  
39 Gly Glu Asn Leu Leu Lys Lys Pro Val Ser Glu Asp Asn Pro Glu Thr  
40 145 150 155 160  
41  
42 Tyr Glu Glu Ala Leu Lys Arg Phe Ala Lys Leu Leu Ser Asp Arg Lys  
43 165 170 175  
44  
45 Lys Leu Arg Ala Asn Lys Ala Ser Tyr Gly Pro Gly Gln Leu Gly Glu  
46 180 185 190  
47  
48 Met Val Thr Val Leu Ser Ile Asp Gly Gly Gly Ile Arg Gly Ile Ile  
49 195 200 205  
50  
51 Pro Ala Thr Ile Leu Glu Phe Leu Glu Gly Gln Leu Gln Glu Met Asp  
52 210 215 220  
53  
54 Asn Asn Ala Asp Ala Arg Leu Ala Asp Tyr Phe Asp Val Ile Gly Gly  
55 225 230 235 240  
56  
57 Thr Ser Thr Gly Leu Leu Thr Ala Met Ile Ser Thr Pro Asn Glu

1	245	250	255
2			
3	Asn Asn Arg Pro Phe Ala Ala Ala Lys Glu Ile Val Pro Phe Tyr Phe		
4	260	265	270
5			
6	Glu His Gly Pro Gln Ile Phe Asn Pro Ser Gly Gln Ile Leu Gly Pro		
7	275	280	285
8			
9	Lys Tyr Asp Gly Lys Tyr Leu Met Gln Val Leu Gln Glu Lys Leu Gly		
10	290	295	300
11			
12	Glu Thr Arg Val His Gln Ala Leu Thr Glu Val Val Ile Ser Ser Phe		
13	305	310	315
14			
15	Asp Ile Lys Thr Asn Lys Pro Val Ile Phe Thr Lys Ser Asn Leu Ala		
16	325	330	335
17			
18	Asn Ser Pro Glu Leu Asp Ala Lys Met Tyr Asp Ile Ser Tyr Ser Thr		
19	340	345	350
20			
21	Ala Ala Ala Pro Thr Tyr Phe Pro Pro His Tyr Phe Val Thr		
22	355	360	365
23			
24	<210> 252		
25	<211> 55		
26	<212> DNA		
27	<213> Synthetic		
28			
29	<400> 252		
30	ggagctcgag aaaagagagg ctgaagctag ttattccaca gcagcagctc caaca		55
31			
32	<210> 253		
33	<211> 39		
34	<212> DNA		
35	<213> Synthetic		
36			
37	<400> 253		
38	ggtctagagg aattctcatt atatgtcata catcttagc		39
39			
40	<210> 254		
41	<211> 1128		
42	<212> DNA		
43	<213> Synthetic		
44			
45	<400> 254		
46	tcgagaaaag agaggctgaa gctagttatt ccacagcgc agctccaaca tattttcctc		60
47			
48	cgcattactt tggtaataat actagtaatg gagatgaata tgagttcaat cttgttgatg		120
49			
50	gtgctgtgc tactgttgct gatccggcgt tattatccat tagcgttgc acgagacttg		180
51			
52	cacaaaagga tccagcattt gcttcaatta ggtcattgaa ttacaaaaaa atgctgttgc		240
53			
54	tctcattagg cactggcact acttcagagt ttgataaaac atatacagca aaagaggcag		300
55			
56	ctacctggac tgctgtacat tggatgttag ttatacagaa aatgactgtat gcagcaagg		360
57			

```

1 cttatcatgac tgattattac ctttctactg ctttcaagc tcttgattca aaaaacaatt 420
2 acctcagggt tcaagaaaat gcattaaacag gcacaactac tgaaatggat gatgcttctg 480
3 aggctaataat ggaatttata gtacaagttg gtgaaaactt attgaagaaa ccagttccg 540
4 aagacaatcc tgaaacctat gaggaagctc taaagaggtt tgcaaaatct ctctctgata 600
5 ggaagaaaact ccgagcaaac aaagcttctt atggaccagg acagttggga gaaatggta 660
6 ctgttcttag tattgtatgga ggttggatatta gagggtatcat tccggcttacc atttcgaat 720
7 ttcttggagg aacaacttcg gaaatggaca ataatgcaga tgcaagactt gcagattact 780
8 ttgatgtat tggaggaaca agtacaggag gtttattgtac tgctatgata agtactccaa 840
9 atgaaaacaa tcgacccttt gctgctgcca aagaaattgt accttttac ttcaacatg 900
10 gccctcagat tttaatctt agtggtaaaa tttagggccc aaaatatgtat gaaaaatata 960
11 ttatgtcaat tcttcaagaa aaacttgggaa aactcgtgt gcatcaagct ttgacagaag 1020
12 ttgtcatctc aagcttgcg atcaaaaacaa ataagccagt aatattcaat aagtcaaaatt 1080
13 tagcaaaactc tccagaattt gatgctaaga tttatgacat ataatgaa 1128
14
15 <210> 255
16 <211> 366
17 <212> PRT
18 <213> Synthetic
19
20 <400> 255
21
22 Ser Tyr Ser Thr Ala Ala Ala Pro Thr Tyr Phe Pro Pro His Tyr Phe
23 1 5 10 15
24
25 Val Thr Asn Thr Ser Asn Gly Asp Glu Tyr Glu Phe Asn Leu Val Asp
26 20 25 30
27
28 Gly Ala Val Ala Thr Val Ala Asp Pro Ala Leu Leu Ser Ile Ser Val
29 35 40 45
30
31 Ala Thr Arg Leu Ala Gln Lys Asp Pro Ala Phe Ala Ser Ile Arg Ser
32 50 55 60
33
34 Leu Asn Tyr Lys Lys Met Leu Leu Leu Ser Leu Gly Thr Gly Thr Thr
35 65 70 75 80
36
37 Ser Glu Phe Asp Lys Thr Tyr Thr Ala Lys Glu Ala Ala Thr Trp Thr
38 85 90 95
39
40 Ala Val His Trp Met Leu Val Ile Gln Lys Met Thr Asp Ala Ala Ser
41 100 105 110
42
43 Ser Tyr Met Thr Asp Tyr Tyr Leu Ser Thr Ala Phe Gln Ala Leu Asp
44 115 120 125
45
46
47
48
49
50
51
52
53
54
55
56
57

```

1    Ser Lys Asn Asn Tyr Leu Arg Val Gln Glu Asn Ala Leu Thr Gly Thr  
2                    130                    135                    140  
3  
4    Thr Thr Glu Met Asp Asp Ala Ser Glu Ala Asn Met Glu Leu Leu Val  
5                    145                    150                    155                    160  
6  
7    Gln Val Gly Glu Asn Leu Leu Lys Lys Pro Val Ser Glu Asp Asn Pro  
8                    165                    170                    175  
9  
10   Glu Thr Tyr Glu Glu Ala Leu Lys Arg Phe Ala Lys Leu Leu Ser Asp  
11                    180                    185                    190  
12  
13   Arg Lys Lys Leu Arg Ala Asn Lys Ala Ser Tyr Gly Pro Gly Gln Leu  
14                    195                    200                    205  
15  
16   Gly Glu Met Val Thr Val Leu Ser Ile Asp Gly Gly Gly Ile Arg Gly  
17                    210                    215                    220  
18  
19   Ile Ile Pro Ala Thr Ile Leu Glu Phe Leu Glu Gly Gln Leu Gln Glu  
20                    225                    230                    235                    240  
21  
22   Met Asp Asn Asn Ala Asp Ala Arg Leu Ala Asp Tyr Phe Asp Val Ile  
23                    245                    250                    255  
24  
25   Gly Gly Thr Ser Thr Gly Gly Leu Leu Thr Ala Met Ile Ser Thr Pro  
26                    260                    265                    270  
27  
28   Asn Glu Asn Asn Arg Pro Phe Ala Ala Ala Lys Glu Ile Val Pro Phe  
29                    275                    280                    285  
30  
31   Tyr Phe Glu His Gly Pro Gln Ile Phe Asn Pro Ser Gly Gln Ile Leu  
32                    290                    295                    300  
33  
34   Gly Pro Lys Tyr Asp Gly Lys Tyr Leu Met Gln Val Leu Gln Glu Lys  
35                    305                    310                    315                    320  
36  
37   Leu Gly Glu Thr Arg Val His Gln Ala Leu Thr Glu Val Val Ile Ser  
38                    325                    330                    335  
39  
40   Ser Phe Asp Ile Lys Thr Asn Lys Pro Val Ile Phe Thr Lys Ser Asn  
41                    340                    345                    350  
42  
43   Leu Ala Asn Ser Pro Glu Leu Asp Ala Lys Met Tyr Asp Ile  
44                    355                    360                    365  
45  
46   <210> 256  
47   <211> 55  
48   <212> DNA  
49   <213> Synthetic  
50  
51   <400> 256  
52   ggagctcgag aaaagagagg ctgaagctac atatacagca aaagaggcag ctacc        55  
53  
54   <210> 257  
55   <211> 39  
56   <212> DNA  
57   <213> Synthetic

```

1 <400> 257
2 ggtctagagg aatttcatt atttatcaa ctctgaagt
3 <210> 258
4 <211> 1128
5 <212> DNA
6 <213> Synthetic
7
8 <400> 258
9 tcgagaaaag agaggtgaa gctacatata cagcaaaaga ggcagctacc tggactgctg 39
10 tacattggat gttagttt cagaaaatg ctgtatgcag aagttttac atgactgatt 60
11 attaccttc tactgtttt caagctttt attcaaaaaa caattacctc agggttcaag 120
12 aaaaatgcatt aacaggcaca actactgaaa tggatgtgc ttctgaggct aatatggaa 180
13 tattagtaca agttgtgaa aacttattga agaaaccagt ttccgaagac aatccgtt 240
14 cctatgagga agctctaaag aggttgaa aattgtctc ttagatggaa aactccgt 300
15 caaacaaggc ttcttatggc ccaggacagt tggagaaat ggtgtactttt cttatgtt 360
16 atggagggtgg aatttagaggg atcattccgg ctaccattct cgaatttctt gaaggacaac 420
17 ttccggaaat ggacaataat gcagatgcga gacttgcaga ttactttgtt gtaattggag 480
18 gaaacaaggta actagggttta ttgtactgtt tgataagtac tccaaatgaa aacaatcgac 540
19 ccttgcgtgc tgccaaagaa attgtacattt ttacttgcg acatggccct cagattttt 600
20 atccatgttgg tcaaaattttt ggcccaaaaat atgtggaaa atatcttgcg caatgtttt 660
21 aaaaaaaact tggagaaact cgtgtgcattt aagcttgcg agaagttgtc atctcaagct 720
22 ttgacatcaa aacaaataa ccagtaatat tcaactaagtccaaatggaa aactctccag 780
23 aatttggatgc taagatgtat gacataaggat attccacagc agcagctcca acatattttt 840
24 ctccgcattt cttgttactt aataactgtt atggagatgtt atatggatgc aatcttgcg 900
25 atttggatgtt gttttttttt gttttttttt cttttttttt cttttttttt cttttttttt 960
26 ttgcacaaaaa ggttcccgat ttgtttttttt cttttttttt cttttttttt cttttttttt 1020
27 tttttttttt cttttttttt cttttttttt cttttttttt cttttttttt cttttttttt 1080
28 tttttttttt cttttttttt cttttttttt cttttttttt cttttttttt cttttttttt 1128
29 <210> 259
30 <211> 366
31 <212> PRT
32 <213> Synthetic
33
34 <400> 259
35 Thr Tyr Thr Ala Lys Glu Ala Ala Thr Trp Thr Ala Val His Trp Met
36 1 5 10 15
37

```

1 Leu Val Ile Gln Lys Met Thr Asp Ala Ala Ser Ser Tyr Met Thr Asp  
2 20 25 30  
3  
4 Tyr Tyr Leu Ser Thr Ala Phe Gln Ala Leu Asp Ser Lys Asn Asn Tyr  
5 35 40 45  
6  
7 Leu Arg Val Gln Glu Asn Ala Leu Thr Gly Thr Thr Glu Met Asp  
8 50 55 60  
9  
10 Asp Ala Ser Glu Ala Asn Met Glu Leu Leu Val Gln Val Gly Glu Asn  
11 65 70 75 80  
12  
13 Leu Leu Lys Lys Pro Val Ser Glu Asp Asn Pro Glu Thr Tyr Glu Glu  
14 85 90 95  
15  
16 Ala Leu Lys Arg Phe Ala Lys Leu Leu Ser Asp Arg Lys Lys Leu Arg  
17 100 105 110  
18  
19 Ser Asn Lys Ala Ser Tyr Gly Pro Gly Gln Leu Gly Glu Met Val Thr  
20 115 120 125  
21  
22 Val Leu Ser Ile Asp Gly Gly Ile Arg Gly Ile Ile Pro Ala Thr  
23 130 135 140  
24  
25 Ile Leu Glu Phe Leu Glu Gly Gln Leu Gln Glu Met Asp Asn Ala  
26 145 150 155 160  
27  
28 Asp Ala Arg Leu Ala Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser Thr  
29 165 170 175  
30  
31 Gly Gly Leu Leu Thr Ala Met Ile Ser Thr Pro Asn Glu Asn Asn Arg  
32 180 185 190  
33  
34 Pro Phe Ala Ala Ala Lys Glu Ile Val Pro Phe Tyr Phe Glu His Gly  
35 195 200 205  
36  
37 Pro Gln Ile Phe Asn Pro Ser Gly Gln Ile Leu Gly Pro Lys Tyr Asp  
38 210 215 220  
39  
40 Gly Lys Tyr Leu Met Gln Val Leu Gln Glu Lys Leu Gly Glu Thr Arg  
41 225 230 235 240  
42  
43 Val His Gln Ala Leu Thr Glu Val Val Ile Ser Ser Phe Asp Ile Lys  
44 245 250 255  
45  
46 Thr Asn Lys Pro Val Ile Phe Thr Lys Ser Asn Leu Ala Asn Ser Pro  
47 260 265 270  
48  
49 Glu Leu Asp Ala Lys Met Tyr Asp Ile Ser Tyr Ser Thr Ala Ala Ala  
50 275 280 285  
51  
52 Pro Thr Tyr Phe Pro Pro His Tyr Phe Val Thr Asn Thr Ser Asn Gly  
53 290 295 300  
54  
55 Asp Glu Tyr Glu Phe Asn Leu Val Asp Gly Ala Val Ala Thr Val Ala  
56 305 310 315 320  
57

1  
 2 Asp Pro Ala Leu Leu Ser Ile Ser Val Ala Thr Arg Leu Ala Gln Lys  
 3 325 330 335  
 4  
 5 Asp Pro Ala Phe Ala Ser Ile Arg Ser Leu Asn Tyr Lys Lys Met Leu  
 6 340 345 350  
 7  
 8 Leu Leu Ser Leu Gly Thr Gly Thr Ser Glu Phe Asp Lys  
 9 355 360 365  
 10  
 11 <210> 260  
 12 <211> 55  
 13 <212> DNA  
 14 <213> Synthetic  
 15  
 16 <400> 260  
 17 ggagctcgag aaaagagagg ctgaagctaa tgcatataa ggcacaacta ctgaa 55  
 18  
 19 <210> 261  
 20 <211> 39  
 21 <212> DNA  
 22 <213> Synthetic  
 23  
 24 <400> 261  
 25 ggtcttagagg aattctcatt attcttgaac cctgaggtta 39  
 26  
 27 <210> 262  
 28 <211> 1128  
 29 <212> DNA  
 30 <213> Synthetic  
 31  
 32 <400> 262  
 33 tcgagaaaaag agaggctgaa gctaatgcata taacaggcac aactactgaa atggatgtat 60  
 34  
 35 cttctgaggc taatatggaa ttatttagtac aagtggtaa aaacttattt aagaaaccag 120  
 36  
 37 tttccgaaga caatctcgaa acctatgagg aagctctaa gaggttgcata aatttgcct 180  
 38  
 39 ctgataggaa gaaactccga gcaaaacaaag cttttatgg accaggacag ttgggagaaa 240  
 40  
 41 tggtgactgt tcttagtatt gatggaggtt gaatttagagg gatcattccg gctaccattc 300  
 42  
 43 tcgaaatttct tgaaggacaa cttcaggaaa tggacataaa tgcatatgcata agacttgcag 360  
 44  
 45 attacttgcata tggatattggaa ggaacaagttt caggatgtt attgtactgtt atgataagtt 420  
 46  
 47 ctccaaatgcata aacaatcgac cccttgcgtt ctggccaaaga aatttgatctt ttttacttcg 480  
 48  
 49 aacatggccc tcagatttt aatccatgtt gtcaaaatttt aggccaaaaa tatgtatggaa 540  
 50  
 51 aatatcttat gcaaggctttt caagaaaaac ttggagaaac tcgtgtgcata caagcttgcata 600  
 52  
 53 cagaagttgtt cactctcaagc ttgacatca aaacaaataa gccagatataa ttcaactaagt 660  
 54  
 55 caaattttgcata aacatctccaa gaattggatgtt ctaagatgtt tgacataagt tattccacat 720  
 56  
 57 cagcagctccaa acatatttt cttccgcattt actttgttac taatactgtt aatggatgtt 780

1 aatatgagtt caaatcttggtt gatggcgctg ttgctactgt tgctgatccg gcgttattat 840  
 2 ccattagcggt tgcacacgaga cttgcacaaa aggatccagc atttgcattca attaggtcat 900  
 3 tgaattacaa aaaaatgcgt ttgctctcat taggcactgg cactacttca gagtttgata 960  
 4 aaacatatac agcaaaaagag gcagctacgtt ggactgtgtt acatggatg ttagttatac 1020  
 5 agaaaatgac tgcacacgaga cttgcacaaa aggatccagc atttgcattca attaggtcat 1080  
 6 aagctcttga ttcaaaaaac aattacctca gggtcaaga ataatgag 1128  
 7  
 8 <210> 263  
 9 <211> 366  
 10 <212> PRT  
 11 <213> Synthetic  
 12  
 13 <400> 263  
 14  
 15 Asn Ala Leu Thr Gly Thr Thr Glu Met Asp Asp Ala Ser Glu Ala 15  
 16 1 5 10 15  
 17 Asn Met Glu Leu Leu Val Gln Val Gly Glu Asn Leu Leu Lys Lys Pro  
 18 20 25 30  
 19 Val Ser Glu Asp Asn Pro Glu Thr Tyr Glu Glu Ala Leu Lys Arg Phe  
 20 35 40 45  
 21 Ala Lys Leu Leu Ser Asp Arg Lys Lys Leu Arg Ala Asn Lys Ala Ser  
 22 50 55 60  
 23 Tyr Gly Pro Gly Gln Leu Gly Glu Met Val Thr Val Leu Ser Ile Asp  
 24 65 70 75 80  
 25 Gly Gly Gly Ile Arg Gly Ile Ile Pro Ala Thr Ile Leu Glu Phe Leu  
 26 85 90 95  
 27 Glu Gly Gln Leu Gln Glu Met Asp Asn Ala Asp Ala Arg Leu Ala  
 28 100 105 110  
 29 Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser Thr Gly Gly Leu Leu Thr  
 30 115 120 125  
 31 Ala Met Ile Ser Thr Pro Asn Glu Asn Asn Arg Pro Phe Ala Ala Ala  
 32 130 135 140  
 33 Lys Glu Ile Val Pro Phe Tyr Phe Glu His Gly Pro Gln Ile Phe Asn  
 34 145 150 155 160  
 35 Pro Ser Gly Gln Ile Leu Gly Pro Lys Tyr Asp Gly Lys Tyr Leu Met  
 36 165 170 175  
 37 Gln Val Leu Gln Glu Lys Leu Gly Glu Thr Arg Val His Gln Ala Leu  
 38 180 185 190  
 39 Thr Glu Val Val Ile Ser Ser Phe Asp Ile Lys Thr Asn Lys Pro Val

1	195	200	205	
2				
3	Ile Phe Thr Lys Ser Asn Leu Ala Asn Ser Pro Glu Leu Asp Ala Lys			
4	210	215	220	
5				
6	Met Tyr Asp Ile Ser Tyr Ser Thr Ala Ala Pro Thr Tyr Phe Pro			
7	225	230	235	240
8				
9	Pro His Tyr Phe Val Thr Asn Thr Ser Asn Gly Asp Glu Tyr Glu Phe			
10	245	250	255	
11				
12	Asn Leu Val Asp Gly Ala Val Ala Thr Val Ala Asp Pro Ala Leu Leu			
13	260	265	270	
14				
15	Ser Ile Ser Val Ala Thr Arg Leu Ala Gln Lys Asp Pro Ala Phe Ala			
16	275	280	285	
17				
18	Ser Ile Arg Ser Leu Asn Tyr Lys Lys Met Leu Leu Leu Ser Leu Gly			
19	290	295	300	
20				
21	Thr Gly Thr Thr Ser Glu Phe Asp Lys Thr Tyr Thr Ala Lys Glu Ala			
22	305	310	315	320
23				
24	Ala Thr Trp Thr Ala Val His Trp Met Leu Val Ile Gln Lys Met Thr			
25	325	330	335	
26				
27	Asp Ala Ala Ser Ser Tyr Met Thr Asp Tyr Tyr Leu Ser Thr Ala Phe			
28	340	345	350	
29				
30	Gln Ala Leu Asp Ser Lys Asn Asn Tyr Leu Arg Val Gln Glu			
31	355	360	365	
32				
33	<210> 264			
34	<211> 1158			
35	<212> DNA			
36	<213> Synthetic			
37				
38	<400> 264			
39	atggccacca ccaagagctt cctcatcctg atcttcatga tcctggccac caccagcagc	60		
40				
41	actttcgccc agctcgccga gatgggtgacc gtgtctcca tcgacggcggt tggcatcagg	120		
42				
43	ggcatcatcc cggccacat cctggagttc ctggaggggc aactccaggaa gatggacaac	180		
44				
45	aacqccgacg cccgcctggc cgactacttc gacgtgtatcg gtggcaccag caccggcggt	240		
46				
47	ctcctgaccc ccatgatctc cactccgaac gagaacaacc gccccttcgc cgctgcgaag	300		
48				
49	gagatcgtcc cgttctactt cgaacacggc cctcagattt tcaacccctc gggtaaatc	360		
50				
51	ctggggccca agtacgacgg caagtacattt atgcaagtgc ttccaggagaa gctggcgag	420		
52				
53	actagggtgc accaggcgct gaccggaggc gtcatctcca gttcgacat caagaccaac	480		
54				
55	aagccagtc tcttccacaa gtccacacgtt gccaacagcc cgaggctgga cgcttaagatg	540		
56				
57	tacgacatct cctactccac tgctgcccgtt cccacgtact tccctccgcatactcgatc	600		

1 accaacaacca gcaacggcga cgagttacgag ttcaacccctt tgacgggtgc ggtggctacg 660  
 2 gttggcgacc cggcgctctt gtcacatcagc gtcgcccacgc gcctggccca gaaggatcca 720  
 3 gccttcgcta gcttagggag cctcaactac aagaagatgc tgctgctcag cctggcact 780  
 4 ggacacacctt ccgagttcga caagacacctt actgccaagg aggcccgtac ctggaccgc 840  
 5 gtccatttga tgctggtcat ccagaagatg acggacgccc cttccagctt catgaccgac 900  
 6 tactacccctt ccacttcggtt ccaggcgctt gacttcaaga acaactaccc cctgttccag 960  
 7 gagaatgccc tcactggcac cacgaccggat atggacgatg cttccggggc caacatggag 1020  
 8 ctgctcgccc aggtgggtga gaaaccttta aagaagcccg tctccgaaaga caatcccgag 1080  
 9 acctatgagg aagcgctcaa ggcgtttgcc aagctgtctt ctgtataggaa gaaactccgc 1140  
 10 gctacaagg ccagctac 1158  
 11  
 12 <210> 265  
 13 <211> 386  
 14 <212> PRT  
 15 <213> Synthetic  
 16  
 17 <400> 265  
 18  
 19 Met Ala Thr Thr Lys Ser Phe Leu Ile Leu Ile Phe Met Ile Leu Ala  
 20 1 5 10 15  
 21  
 22 Thr Thr Ser Ser Thr Phe Ala Gln Leu Gly Glu Met Val Thr Val Leu  
 23 20 25 30  
 24  
 25 Ser Ile Asp Gly Gly Ile Arg Gly Ile Ile Pro Ala Thr Ile Leu  
 26 35 40 45  
 27  
 28 Glu Phe Leu Glu Gly Gln Leu Gln Glu Met Asp Asn Asn Ala Asp Ala  
 29 50 55 60  
 30  
 31 Arg Leu Ala Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser Thr Gly Gly  
 32 65 70 75 80  
 33  
 34 Leu Leu Thr Ala Met Ile Ser Thr Pro Asn Glu Asn Asn Arg Pro Phe  
 35 85 90 95  
 36  
 37 Ala Ala Ala Lys Glu Ile Val Pro Phe Tyr Phe Glu His Gly Pro Gln  
 38 100 105 110  
 39  
 40 Ile Phe Asn Pro Ser Gly Gln Ile Leu Gly Pro Lys Tyr Asp Gly Lys  
 41 115 120 125  
 42  
 43 Tyr Leu Met Gln Val Leu Gln Glu Lys Leu Gly Glu Thr Arg Val His  
 44 130 135 140  
 45  
 46 Gln Ala Leu Thr Glu Val Val Ile Ser Ser Phe Asp Ile Lys Thr Asn  
 47 145 150 155  
 48 160

1 Lys Pro Val Ile Phe Thr Lys Ser Asn Leu Ala Asn Ser Pro Glu Leu  
 2 165 170 175  
 3  
 4 Asp Ala Lys Met Tyr Asp Ile Ser Tyr Ser Thr Ala Ala Ala Pro Thr  
 5 180 185 190  
 6  
 7 Tyr Phe Pro Pro His Tyr Phe Val Thr Asn Thr Ser Asn Gly Asp Glu  
 8 195 200 205  
 9  
 10 Tyr Glu Phe Asn Leu Val Asp Gly Ala Val Ala Thr Val Ala Asp Pro  
 11 210 215 220  
 12  
 13 Ala Leu Leu Ser Ile Ser Val Ala Thr Arg Leu Ala Gln Lys Asp Pro  
 14 225 230 235 240  
 15  
 16 Ala Phe Ala Ser Ile Arg Ser Leu Asn Tyr Lys Lys Met Leu Leu Leu  
 17 245 250 255  
 18  
 19 Ser Leu Gly Thr Gly Thr Thr Ser Glu Phe Asp Lys Thr Tyr Thr Ala  
 20 260 265 270  
 21  
 22 Lys Glu Ala Ala Thr Trp Thr Ala Val His Trp Met Leu Val Ile Gln  
 23 275 280 285  
 24  
 25 Lys Met Thr Asp Ala Ala Ser Ser Tyr Met Thr Asp Tyr Tyr Leu Ser  
 26 290 295 300  
 27  
 28 Thr Ala Phe Gln Ala Leu Asp Ser Lys Asn Asn Tyr Leu Arg Val Gln  
 29 305 310 315 320  
 30  
 31 Glu Asn Ala Leu Thr Gly Thr Thr Glu Met Asp Asp Ala Ser Glu  
 32 325 330 335  
 33  
 34 Ala Asn Met Glu Leu Leu Val Gln Val Gly Glu Asn Leu Leu Lys Lys  
 35 340 345 350  
 36  
 37 Pro Val Ser Glu Asp Asn Pro Glu Thr Tyr Glu Glu Ala Leu Lys Arg  
 38 355 360 365  
 39  
 40 Phe Ala Lys Leu Leu Ser Asp Arg Lys Lys Leu Arg Ala Asn Lys Ala  
 41 370 375 380  
 42  
 43 Ser Tyr  
 44 385  
 45  
 46  
 47 <210> 266  
 48 <211> 55  
 49 <212> DNA  
 50 <213> Synthetic  
 51  
 52 <400> 266  
 53 ggagctcgag aaaagagagg ctgaagctag cctcaactac aagaagatgc tgctg 55  
 54  
 55 <210> 267  
 56 <211> 42  
 57 <212> DNA

```

1 <213> Synthetic
2
3 <400> 267
4 gccgagctgt cctggccgt agctggccgt gtttagcgcgg ag
5
6 <210> 268
7 <211> 36
8 <212> DNA
9 <213> Synthetic
10
11 <400> 268
12 gccagctacg gaccaggaca gctcgccgag atgggt 36
13
14 <210> 269
15 <211> 39
16 <212> DNA
17 <213> Synthetic
18
19 <400> 269
20 ggtctagagg aattctcatt acctaattgtc agcgaaggc 39
21
22 <210> 270
23 <211> 1167
24 <212> DNA
25 <213> Synthetic
26
27 <400> 270
28 atggccacca ccaagagctt cctcattccgt atcttcatgt tcctggccac caccaggcage 60
29
30 accttcgcca gcctcaacta caagaagatctgctgtcgtca gcctggccac tggcacgacc 120
31
32 tccgaggatcc acaagagacca cactgccaag gaggccgcata cctggaccgc cgtccattgg 180
33
34 atgctggtca tccagaagat gacggacgcc gcttccagct acatgaccga ctactacctc 240
35
36 tccacttcgtc tccaggcgtc tgacttcaag aacaactacc tccgtgttca ggagaatggc 300
37
38 ctcacttgca ccacgaccga gatggacgt gcttccggagg ccaacatggc gctgtctgtc 360
39
40 caggtgggtg agaacctctt gaagaagcccc gcttccggagg acaatccccgac gacctatgg 420
41
42 gaagcgctca aegcgcttgc caagctgtc tctgtatggaa agaaactccgc cgcttaacaag 480
43
44 gccagctacg gaccaggaca gctcgccgag atggtgaccgt tgctctccat cgacggcggt 540
45
46 ggcacatggc gcatcatccc ggccaccatc ctggagttcc tggaggggccaa actccaggag 600
47
48 atggacaaaca acgcccacgc ecgcctggcc gactacttcg acgtgtatggc tggaccaggc 660
49
50 accggcggtc tcctgaccgc catgtatccc actccggaaacg agaacaacccg ccccttcgc 720
51
52 gctcgcaagg agatcggtcc gttctacttc gaacacggcc ctcagatttt caacccctcg 780
53
54 ggtcaaatcc tggggcccaa gtacgacggc aagtacctta tgcaagtgtc tcaggagaag 840
55
56 ctggggcaga ctagggtgtca ccaggccgtc accggaggatcg tcatctccag ctgcacatc 900
57

```

1 aagaccaaca agccagtcat cttcaccaag tccaaacctgg ccaacagccc ggagctggac 960  
 2 gctaaagatgt acgacatctc ctactccact gctgccgctc ccacgtactt ccctccgcac 1020  
 3 tacttcgtca ccaacaccag caacggcgac gagtacgagt tcaaccttgt tgacggtgcg 1080  
 4 gttggctacgg tggcgaccc ggcgcctctg tccatcagcg tcgcccccg cctggccag 1140  
 5 aaggatccag ccttcgtctg cattagg 1167  
 6  
 7 <210> 271  
 8 <211> 389  
 9 <212> PRT  
 10 <213> Synthetic  
 11 <400> 271  
 12  
 13 Met Ala Thr Thr Lys Ser Phe Leu Ile Leu Ile Phe Met Ile Leu Ala 18  
 14 5 10 15  
 15  
 16 Thr Thr Ser Ser Thr Phe Ala Ser Leu Asn Tyr Lys Lys Met Leu Leu 21  
 17 20 25 30  
 18 Leu Ser Leu Gly Thr Gly Thr Ser Glu Phe Asp Lys Thr Tyr Thr 24  
 19 35 40 45  
 20  
 21 Ala Lys Glu Ala Ala Thr Trp Thr Ala Val His Trp Met Leu Val Ile 27  
 22 50 55 60  
 23  
 24 Gln Lys Met Thr Asp Ala Ala Ser Ser Tyr Met Thr Asp Tyr Tyr Leu 28  
 25 65 70 75 80  
 26  
 27 Ser Thr Ala Phe Gln Ala Leu Asp Ser Lys Asn Asn Tyr Leu Arg Val 33  
 28 85 90 95  
 29  
 30 Gln Glu Asn Ala Leu Thr Gly Thr Thr Glu Met Asp Asp Ala Ser 36  
 31 100 105 110  
 32  
 33 Glu Ala Asn Met Glu Leu Leu Val Gln Val Gly Glu Asn Leu Leu Lys 39  
 34 115 120 125  
 35  
 36 Lys Pro Val Ser Glu Asp Asn Pro Glu Thr Tyr Glu Glu Ala Leu Lys 42  
 37 130 135 140  
 38  
 39 Arg Phe Ala Lys Leu Leu Ser Asp Arg Lys Lys Leu Arg Ala Asn Lys 45  
 40 145 150 155 160  
 41  
 42 Ala Ser Tyr Gly Pro Gly Gln Leu Gly Glu Met Val Thr Val Leu Ser 48  
 43 165 170 175  
 44  
 45 Ile Asp Gly Gly Gly Ile Arg Gly Ile Ile Pro Ala Thr Ile Leu Glu 51  
 46 180 185 190  
 47  
 48 Phe Leu Glu Gly Gln Leu Gln Glu Met Asp Asn Asn Ala Asp Ala Arg 54  
 49 195 200 205  
 50  
 51 Leu Ala Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser Thr Gly Gly Leu 57

1	210	215	220	
2				
3	Leu Thr Ala Met Ile Ser Thr Pro Asn Glu Asn Asn Arg Pro Phe Ala			
4	225	230	235	240
5				
6	Ala Ala Lys Glu Ile Val Pro Phe Tyr Phe Glu His Gly Pro Gln Ile			
7	245	250		255
8				
9	Phe Asn Pro Ser Gly Gln Ile Leu Gly Pro Lys Tyr Asp Gly Lys Tyr			
10	260	265		270
11				
12	Leu Met Gln Val Leu Gln Glu Lys Leu Gly Glu Thr Arg Val His Gln			
13	275	280		285
14				
15	Ala Leu Thr Glu Val Val Ile Ser Ser Phe Asp Ile Lys Thr Asn Lys			
16	290	295		300
17				
18	Pro Val Ile Phe Thr Lys Ser Asn Leu Ala Asn Ser Pro Glu Leu Asp			
19	305	310	315	320
20				
21	Ala Lys Met Tyr Asp Ile Ser Tyr Ser Thr Ala Ala Ala Pro Thr Tyr			
22	325	330		335
23				
24	Phe Pro Pro His Tyr Phe Val Thr Asn Thr Ser Asn Gly Asp Glu Tyr			
25	340	345		350
26				
27	Glu Phe Asn Leu Val Asp Gly Ala Val Ala Thr Val Ala Asp Pro Ala			
28	355	360		365
29				
30	Leu Leu Ser Ile Ser Val Ala Thr Arg Leu Ala Gln Lys Asp Pro Ala			
31	370	375		380
32				
33	Phe Ala Ser Ile Arg			
34	385			
35				
36	<210> 272			
37	<211> 55			
38	<212> DNA			
39	<213> Synthetic			
40				
41	<400> 272			
42	ggagctcgag aaaaagagagg ctgaagctac tgccaaaggag gccgctacct ggacc			55
43				
44	<210> 273			
45	<211> 39			
46	<212> DNA			
47	<213> Synthetic			
48				
49	<400> 273			
50	ggtctagagg aattctcatt acttgtcga ctcggaggt			39
51				
52	<210> 274			
53	<211> 1167			
54	<212> DNA			
55	<213> Synthetic			
56				
57	<400> 274			

1	atggccacca ccaagagctt cctcatctg atcttcatga tcctggccac caccagcgc	60
2	actttcgcca cctacactgc caaggaggcc gctacctga ccgcgcctca ttggatgctg	120
3	gtcatccaga agatgacgga cgcgcgttcc agtacatgta ccgactacta cctcttccact	180
4	cggttccagg cgcttgactc caagaacaac tacctccgtt ttcaggagaa tgccctcact	240
5	ggcaccacga ccgagatgga cgatgcctcc gaggccaaca tggagctgct cgtccagg	300
6	ggtgagaacc tcctgaagaa gcccgtctcc gaagacaatc ccgagaccta tgaggaagcg	360
7	ctcaagcgct ttgccaagct gctctctgat aggaagaaac tccgcgttaa caagggcagc	420
8	tacgaccacg gacagctcg gtagatggt accgtgtct ccatcgacgg cgggtggatc	480
9	agggggcatca tccccggccac cattctggat ttccctggagg gccaactcca ggagatggac	540
10	aacaacccgc acggccgcct ggcgactact ttcgacgtga tcgggtggac cagcaccggc	600
11	ggtctcctga ccgcctatgat ctccactccg aacgagaaca accgcgcctt cgcgcgtcg	660
12	aggagatcg tcccggttcta cttcgaacac ggcgcctcaga tttcaaccc ctgggttcaa	720
13	atccctggcc ccaagatcga cggcaagact cttatgcaag tgcttcagga gaagctggc	780
14	gagactaggg tgccacccggc gctgaccgg gtcgtcatct ccagcttgcg catcaagacc	840
15	aacaaggccg tcatcttcac caagtccaaac ctggcaaca gcccggagct ggacgctaa	900
16	atgtacgaca tctctactc cactgtgtcc gtcctccatgt acttcctcc gcactacttc	960
17	gtccaaca ccagcaacccg cgcgcgtact gaggtaacc ttgttgcgg tgccgtggct	1020
18	acgggtggcg accccggcgct cctgtccatc agcgtcgcca cgcgcctggc ccagaaggat	1080
19	ccagccttcg ctgcattag ggcctcaac tacaagaaga tgctgtgct cagcctggc	1140
20	actggcacga cctccggatgg cgacaaag	1167
21	<210> 275	
22	<211> 389	
23	<212> PRT	
24	<213> Synthetic	
25	<400> 275	
26	Met Ala Thr Thr Ser Phe Leu Ile Leu Ile Phe Met Ile Leu Ala	
27	1 5 10 15	
28	Thr Thr Ser Ser Thr Phe Ala Thr Tyr Thr Ala Lys Glu Ala Ala Thr	
29	20 25 30	
30	Trp Thr Ala Val His Trp Met Leu Val Ile Gln Lys Met Thr Asp Ala	
31	35 40 45	
32	Ala Ser Ser Tyr Met Thr Asp Tyr Tyr Leu Ser Thr Ala Phe Gln Ala	

1	50	55	60
2			
3	Leu Asp Ser Lys Asn Asn Tyr Leu Arg Val Gln Glu Asn Ala Leu Thr		
4	65 70	75	80
5			
6	Gly Thr Thr Glu Met Asp Asp Ala Ser Glu Ala Asn Met Glu Leu		
7	85	90	95
8			
9	Leu Val Gln Val Gly Glu Asn Leu Leu Lys Lys Pro Val Ser Glu Asp		
10	100	105	110
11			
12	Asn Pro Glu Thr Tyr Glu Glu Ala Leu Lys Arg Phe Ala Lys Leu Leu		
13	115	120	125
14			
15	Ser Asp Arg Lys Lys Leu Arg Ala Asn Lys Ala Ser Tyr Gly Pro Gly		
16	130 135	140	
17			
18	Gln Leu Gly Glu Met Val Thr Val Leu Ser Ile Asp Gly Gly Ile		
19	145 150	155	160
20			
21	Arg Gly Ile Ile Pro Ala Thr Ile Leu Glu Phe Leu Glu Gly Gln Leu		
22	165	170	175
23			
24	Gln Glu Met Asp Asn Asn Ala Asp Ala Arg Leu Ala Asp Tyr Phe Asp		
25	180	185	190
26			
27	Val Ile Gly Gly Thr Ser Thr Gly Gly Leu Leu Thr Ala Met Ile Ser		
28	195	200	205
29			
30	Thr Pro Asn Glu Asn Asn Arg Pro Phe Ala Ala Ala Lys Glu Ile Val		
31	210	215	220
32			
33	Pro Phe Tyr Phe Glu His Gly Pro Gln Ile Phe Asn Pro Ser Gly Gln		
34	225 230	235	240
35			
36	Ile Leu Gly Pro Lys Tyr Asp Gly Lys Tyr Leu Met Gln Val Leu Gln		
37	245	250	255
38			
39	Glu Lys Leu Gly Glu Thr Arg Val His Gln Ala Leu Thr Glu Val Val		
40	260	265	270
41			
42	Ile Ser Ser Phe Asp Ile Lys Thr Asn Lys Pro Val Ile Phe Thr Lys		
43	275	280	285
44			
45	Ser Asn Leu Ala Asn Ser Pro Glu Leu Asp Ala Lys Met Tyr Asp Ile		
46	290	295	300
47			
48	Ser Tyr Ser Thr Ala Ala Pro Thr Tyr Phe Pro Pro His Tyr Phe		
49	305 310	315	320
50			
51	Val Thr Asn Thr Ser Asn Gly Asp Glu Tyr Glu Phe Asn Leu Val Asp		
52	325	330	335
53			
54	Gly Ala Val Ala Thr Val Ala Asp Pro Ala Leu Leu Ser Ile Ser Val		
55	340	345	350
56			
57	Ala Thr Arg Leu Ala Gln Lys Asp Pro Ala Phe Ala Ser Ile Arg Ser		

3 355 360 365  
 2 Leu Asn Tyr Lys Lys Met Leu Leu Leu Ser Leu Gly Thr Gly Thr Thr  
 3 370 375 380  
 4  
 5  
 6 Ser Glu Phe Asp Lys  
 7 385  
 8  
 9 <210> 276  
 10 <211> 7  
 11 <212> PRT  
 12 <213> Synthetic  
 13  
 14 <400> 276  
 15  
 16 Gly Gly Gly Ser Gly Gly Gly  
 17 1 5  
 18  
 19 <210> 277  
 20 <211> 3  
 21 <212> PRT  
 22 <213> Synthetic  
 23  
 24 <400> 277  
 25  
 26 Gly Pro Gly  
 27 1  
 28  
 29 <210> 278  
 30 <211> 386  
 31 <212> PRT  
 32 <213> Solanum tuberosum  
 33  
 34 <400> 278  
 35  
 36 Met Ala Thr Thr Lys Ser Phe Leu Ile Leu Phe Phe Met Ile Leu Ala  
 37 1 5 10 15  
 38  
 39 Thr Thr Ser Ser Thr Cys Ala Lys Leu Glu Glu Met Val Thr Val Leu  
 40 20 25 30  
 41  
 42 Ser Ile Asp Gly Gly Ile Lys Gly Ile Ile Pro Ala Ile Ile Leu  
 43 35 40 45  
 44  
 45 Glu Phe Leu Glu Gly Gln Leu Gln Glu Val Asp Asn Asn Lys Asp Ala  
 46 50 55 60  
 47  
 48 Arg Leu Ala Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser Thr Gly Gly  
 49 65 70 75 80  
 50  
 51 Leu Leu Thr Ala Met Ile Thr Thr Pro Asn Glu Asn Asn Arg Pro Phe  
 52 85 90 95  
 53  
 54 Ala Ala Ala Lys Asp Ile Val Pro Phe Tyr Phe Glu His Gly Pro His  
 55 100 105 110  
 56  
 57 Ile Phe Asn Tyr Ser Gly Ser Ile Ile Gly Pro Met Tyr Asp Gly Lys

1	115	120	125
2			
3	Tyr Leu Leu Gln Val Leu Gln Glu Lys Leu Gly Glu Thr Arg Val His		
4	130	135	140
5			
6	Gln Ala Leu Thr Glu Val Ala Ile Ser Ser Phe Asp Ile Lys Thr Asn		
7	145	150	155
8			160
9	Lys Pro Val Ile Phe Thr Lys Ser Asn Leu Ala Lys Ser Pro Glu Leu		
10	165	170	175
11			
12	Asp Ala Lys Met Tyr Asp Ile Cys Tyr Ser Thr Ala Ala Ala Pro Ile		
13	180	185	190
14			
15	Tyr Phe Pro Pro His Tyr Phe Ile Thr His Thr Ser Asn Gly Asp Ile		
16	195	200	205
17			
18	Tyr Glu Phe Asn Leu Val Asp Gly Gly Val Ala Thr Val Gly Asp Pro		
19	210	215	220
20			
21	Ala Leu Leu Ser Leu Ser Val Ala Thr Arg Leu Ala Gln Glu Asp Pro		
22	225	230	235
23			240
24	Ala Phe Ser Ser Ile Lys Ser Leu Asp Tyr Lys Gln Met Leu Leu Leu		
25	245	250	255
26			
27	Ser Leu Gly Thr Gly Thr Asn Ser Glu Phe Asp Lys Thr Tyr Thr Ala		
28	260	265	270
29			
30	Gln Glu Ala Ala Lys Trp Gly Pro Leu Arg Trp Met Leu Ala Ile Gln		
31	275	280	285
32			
33	Gln Met Thr Asn Ala Ala Ser Ser Tyr Met Thr Asp Tyr Tyr Ile Ser		
34	290	295	300
35			
36	Thr Val Phe Gln Ala Arg His Ser Gln Asn Asn Tyr Leu Arg Val Gln		
37	305	310	315
38			320
39	Glu Asn Ala Leu Thr Gly Thr Thr Glu Met Asp Asp Ala Ser Glu		
40	325	330	335
41			
42	Ala Asn Met Glu Leu Leu Val Gln Val Gly Glu Thr Leu Leu Lys Lys		
43	340	345	350
44			
45	Pro Val Ser Lys Asp Ser Pro Glu Thr Tyr Glu Glu Ala Leu Lys Arg		
46	355	360	365
47			
48	Phe Ala Lys Leu Leu Ser Asp Arg Lys Lys Leu Arg Ala Asn Lys Ala		
49	370	375	380
50			
51	Ser Tyr		
52	385		
53			
54	<210> 279		
55	<211> 386		
56	<212> PRT		
57	<213> Solanum tuberosum		

1  
2 <400> 279  
3  
4 Met Ala Thr Thr Lys Ser Val Leu Val Leu Phe Phe Met Ile Leu Ala  
5 1 5 10 15  
6  
7 Thr Thr Ser Ser Thr Cys Ala Thr Leu Gly Glu Met Val Thr Val Leu  
8 20 25 30  
9  
10 Ser Ile Asp Gly Gly Ile Lys Gly Ile Ile Pro Ala Thr Ile Leu  
11 35 40 45  
12  
13 Glu Phe Leu Glu Gly Gln Leu Gln Glu Val Asp Asn Asn Lys Asp Ala  
14 50 55 60  
15  
16 Arg Leu Ala Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser Thr Gly Gly  
17 65 70 75 80  
18  
19 Leu Leu Thr Ala Met Ile Thr Thr Pro Asn Glu Asn Asn Arg Pro Phe  
20 85 90 95  
21  
22 Ala Ala Ala Lys Asp Ile Val Pro Phe Tyr Phe Glu His Gly Pro His  
23 100 105 110  
24  
25 Ile Phe Asn Ser Ser Gly Ser Ile Phe Gly Pro Met Tyr Asp Gly Lys  
26 115 120 125  
27  
28 Tyr Phe Leu Gln Val Leu Gln Glu Lys Leu Gly Glu Thr Arg Val His  
29 130 135 140  
30  
31 Gln Ala Leu Thr Glu Val Ala Ile Ser Ser Phe Asp Ile Lys Thr Asn  
32 145 150 155 160  
33  
34 Lys Pro Val Ile Phe Thr Lys Ser Asn Leu Ala Lys Ser Pro Glu Leu  
35 165 170 175  
36  
37 Asp Ala Lys Met Asn Asp Ile Cys Tyr Ser Thr Ala Ala Ala Pro Thr  
38 180 185 190  
39  
40 Tyr Phe Pro Pro His Tyr Phe Val Thr His Thr Ser Asn Gly Asp Lys  
41 195 200 205  
42  
43 Tyr Glu Phe Asn Leu Val Asp Gly Ala Val Ala Thr Val Gly Asp Pro  
44 210 215 220  
45  
46 Ala Leu Leu Ser Leu Ser Val Arg Thr Lys Leu Ala Gln Val Asp Pro  
47 225 230 235 240  
48  
49 Lys Phe Ala Ser Ile Lys Ser Leu Asn Tyr Asn Glu Met Leu Leu Leu  
50 245 250 255  
51  
52 Ser Leu Gly Thr Gly Thr Asn Ser Glu Phe Asp Lys Thr Tyr Thr Ala  
53 260 265 270  
54  
55 Glu Glu Ala Ala Lys Trp Gly Pro Leu Arg Trp Ile Leu Ala Ile Gln  
56 275 280 285  
57

1 Gln Met Thr Asn Ala Ala Ser Ser Tyr Met Thr Asp Tyr Tyr Leu Ser  
 2 290 295 300  
 3  
 4 Thr Val Phe Gln Ala Arg His Ser Gln Asn Asn Tyr Leu Arg Val Gln  
 5 305 310 315 320  
 6  
 7 Glu Asn Ala Leu Thr Gly Thr Thr Glu Met Asp Asp Ala Ser Glu  
 8 325 330 335  
 9  
 10 Ala Asn Met Glu Leu Leu Val Gln Val Gly Glu Lys Leu Leu Lys Lys  
 11 340 345 350  
 12  
 13 Pro Val Ser Lys Asp Ser Pro Glu Thr Tyr Glu Glu Ala Leu Lys Arg  
 14 355 360 365  
 15  
 16 Phe Ala Lys Leu Leu Ser Asp Arg Lys Lys Leu Arg Ala Asn Lys Ala  
 17 370 375 380  
 18  
 19 Ser Tyr  
 20 385  
 21  
 22 <210> 280  
 23 <211> 365  
 24 <212> PRT  
 25 <213> Solanum tuberosum  
 26  
 27 <400> 280  
 28  
 29 Met Ala Leu Glu Glu Met Val Ala Val Leu Ser Ile Asp Gly Gly Gly  
 30 1 5 10 15  
 31  
 32 Ile Lys Gly Ile Ile Pro Gly Thr Ile Leu Glu Phe Leu Glu Gly Gln  
 33 20 25 30  
 34  
 35 Leu Gln Lys Met Asp Asn Asn Ala Asp Ala Arg Leu Ala Asp Tyr Phe  
 36 35 40 45  
 37  
 38 Asp Val Ile Gly Gly Thr Ser Thr Gly Gly Leu Leu Thr Ala Met Ile  
 39 50 55 60  
 40  
 41 Thr Thr Pro Asn Glu Asn Asn Arg Pro Phe Ala Ala Ala Asn Glu Ile  
 42 65 70 75 80  
 43  
 44 Val Pro Phe Tyr Phe Glu His Gly Pro His Ile Phe Asn Ser Arg Tyr  
 45 85 90 95  
 46  
 47 Trp Pro Ile Phe Trp Pro Lys Tyr Asp Gly Lys Tyr Leu Met Gln Val  
 48 100 105 110  
 49  
 50 Leu Gln Glu Lys Leu Gly Glu Thr Arg Val His Gln Ala Leu Thr Glu  
 51 115 120 125  
 52  
 53 Val Ala Ile Ser Ser Phe Asp Ile Lys Thr Asn Lys Pro Val Ile Phe  
 54 130 135 140  
 55  
 56 Thr Lys Ser Asn Leu Ala Lys Ser Pro Glu Leu Asp Ala Lys Thr Tyr  
 57 145 150 155 160

1 Asp Ile Cys Tyr Ser Thr Ala Ala Ala Pro Thr Tyr Phe Pro Pro His  
 2 165 170 175  
 3  
 4 Tyr Phe Ala Thr Asn Thr Ile Asn Gly Asp Lys Tyr Glu Phe Asn Leu  
 5 180 185 190  
 6  
 7 Val Asp Gly Ala Val Ala Thr Val Ala Asp Pro Ala Leu Leu Ser Val  
 8 195 200 205  
 9  
 10 Ser Val Ala Thr Arg Arg Ala Gln Glu Asp Pro Ala Phe Ala Ser Ile  
 11 210 215 220  
 12  
 13 Arg Ser Leu Asn Tyr Lys Lys Met Leu Leu Leu Ser Leu Gly Thr Gly  
 14 225 230 235 240  
 15  
 16 Thr Thr Ser Glu Phe Asp Lys Thr His Thr Ala Glu Glu Thr Ala Lys  
 17 245 250 255  
 18  
 19 Trp Gly Ala Leu Gln Trp Met Leu Val Ile Gln Gln Met Thr Glu Ala  
 20 260 265 270  
 21  
 22 Ala Ser Ser Tyr Met Thr Asp Tyr Tyr Leu Ser Thr Val Phe Gln Asp  
 23 275 280 285  
 24  
 25 Leu His Ser Gln Asn Asn Tyr Leu Arg Val Gln Glu Asn Ala Leu Thr  
 26 290 295 300  
 27  
 28 Gly Thr Thr Thr Lys Ala Asp Asp Ala Ser Glu Ala Asn Met Glu Leu  
 29 305 310 315 320  
 30  
 31 Leu Ala Gln Val Gly Glu Asn Leu Leu Lys Lys Pro Val Ser Lys Asp  
 32 325 330 335  
 33  
 34 Asn Pro Glu Thr Tyr Glu Glu Ala Leu Lys Arg Phe Ala Lys Leu Leu  
 35 340 345 350  
 36  
 37 Ser Asp Arg Lys Lys Leu Arg Ala Asn Lys Ala Ser Tyr  
 38 355 360 365  
 39  
 40 <210> 281  
 41 <211> 364  
 42 <212> PRT  
 43 <213> Solanum tuberosum  
 44  
 45 <400> 281  
 46  
 47 Pro Trp Leu Glu Glu Met Val Thr Val Leu Ser Ile Asp Gly Gly  
 48 1 5 10 15  
 49  
 50 Ile Lys Gly Ile Ile Pro Ala Ile Ile Leu Glu Phe Leu Glu Gly Gln  
 51 20 25 30  
 52  
 53 Leu Gln Glu Val Asp Asn Asn Lys Asp Ala Arg Leu Ala Asp Tyr Phe  
 54 35 40 45  
 55  
 56 Asp Val Ile Gly Gly Thr Ser Thr Gly Leu Leu Thr Ala Met Ile  
 57

1 50 55 60  
2  
3 Thr Thr Pro Asn Glu Asn Asn Arg Pro Phe Ala Ala Ala Lys Asp Ile  
4 65 70 75 80  
5  
6 Val Pro Phe Tyr Phe Glu His Gly Pro His Ile Phe Asn Tyr Ser Gly  
7 85 90 95  
8  
9 Ser Ile Leu Gly Pro Met Tyr Asp Gly Lys Tyr Leu Leu Gln Val Leu  
10 100 105 110  
11  
12 Gln Glu Lys Leu Gly Glu Thr Arg Val His Gln Ala Leu Thr Glu Val  
13 115 120 125  
14  
15 Ala Ile Ser Ser Phe Asp Ile Lys Thr Asn Lys Pro Val Ile Phe Thr  
16 130 135 140  
17  
18 Lys Ser Asn Leu Ala Lys Ser Pro Glu Leu Asp Ala Lys Met Tyr Asp  
19 145 150 155 160  
20  
21 Ile Cys Tyr Ser Thr Ala Ala Ala Pro Ile Tyr Phe Pro Pro His His  
22 165 170 175  
23  
24 Phe Val Thr His Thr Ser Asn Gly Ala Arg Tyr Glu Phe Asn Leu Val  
25 180 185 190  
26  
27 Asp Gly Ala Val Ala Thr Val Gly Asp Pro Ala Leu Leu Ser Leu Ser  
28 195 200 205  
29  
30 Val Ala Thr Arg Leu Ala Gln Glu Asp Pro Ala Phe Ser Ser Ile Lys  
31 210 215 220  
32  
33 Ser Leu Asp Tyr Lys Gln Met Leu Leu Leu Ser Leu Gly Thr Gly Thr  
34 225 230 235 240  
35  
36 Asn Ser Glu Phe Asp Lys Thr Tyr Thr Ala Glu Glu Ala Ala Lys Trp  
37 245 250 255  
38  
39 Gly Pro Leu Arg Trp Met Leu Ala Ile Gln Gln Met Thr Asn Ala Ala  
40 260 265 270  
41  
42 Ser Phe Tyr Met Thr Asp Tyr Tyr Ile Ser Thr Val Phe Gln Ala Arg  
43 275 280 285  
44  
45 His Ser Gln Asn Asn Tyr Leu Arg Val Gln Glu Asn Ala Leu Asn Gly  
46 290 295 300  
47  
48 Thr Thr Thr Glu Met Asp Asp Ala Ser Glu Ala Asn Met Glu Leu Leu  
49 305 310 315 320  
50  
51 Val Gln Val Gly Glu Thr Leu Leu Lys Lys Pro Val Ser Arg Asp Ser  
52 325 330 335  
53  
54 Pro Glu Thr Tyr Glu Glu Ala Leu Lys Arg Phe Ala Lys Leu Leu Ser  
55 340 345 350  
56  
57 Asp Arg Lys Lys Leu Arg Ala Asn Lys Ala Ser Tyr

1	355	360
2		
3	<210> 282	
4	<211> 386	
5	<212> PRT	
6	<213> Solanum tuberosum	
7		
8	<400> 282	
9		
10	Met Ala Thr Thr Lys Ser Phe Leu Ile Leu Phe Phe Met Ile Leu Ala	
11	1 5 10 15	
12		
13	Thr Thr Ser Ser Thr Cys Ala Lys Leu Glu Glu Met Val Thr Val Leu	
14	20 25 30	
15		
16	Ser Ile Asp Gly Gly Ile Lys Gly Ile Ile Pro Ala Ile Ile Leu	
17	35 40 45	
18		
19	Glu Phe Leu Glu Gly Gln Leu Gln Glu Val Asp Asn Asn Lys Asp Ala	
20	50 55 60	
21		
22	Arg Leu Ala Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser Thr Gly Gly	
23	65 70 75 80	
24		
25	Leu Leu Thr Ala Met Ile Thr Thr Pro Asn Glu Asn Asn Arg Pro Phe	
26	85 90 95	
27		
28	Ala Ala Ala Lys Asp Ile Val Pro Phe Tyr Phe Glu His Gly Pro His	
29	100 105 110	
30		
31	Ile Phe Asn Tyr Ser Gly Ser Ile Leu Gly Pro Met Tyr Asp Gly Lys	
32	115 120 125	
33		
34	Tyr Leu Leu Gln Val Leu Gln Glu Lys Leu Gly Glu Thr Arg Val His	
35	130 135 140	
36		
37	Gln Ala Leu Thr Glu Val Ala Ile Ser Ser Phe Asp Ile Lys Thr Asn	
38	145 150 155 160	
39		
40	Lys Pro Val Ile Phe Thr Lys Ser Asn Leu Ala Lys Ser Pro Glu Leu	
41	165 170 175	
42		
43	Asp Ala Lys Met Tyr Asp Ile Cys Tyr Ser Thr Ala Ala Ala Pro Ile	
44	180 185 190	
45		
46	Tyr Phe Pro Pro His His Phe Val Thr His Thr Ser Asn Gly Ala Arg	
47	195 200 205	
48		
49	Tyr Glu Phe Asn Leu Val Asp Gly Ala Val Ala Thr Val Gly Asp Pro	
50	210 215 220	
51		
52	Ala Leu Leu Ser Leu Ser Val Ala Thr Arg Leu Ala Gln Glu Asp Pro	
53	225 230 235 240	
54		
55	Ala Phe Ser Ser Ile Lys Ser Leu Asp Tyr Lys Gln Met Leu Leu Leu	
56	245 250 255	
57		

1    Ser Leu Gly Thr Gly Thr Asn Ser Glu Phe Asp Lys Thr Tyr Thr Ala  
 2                    260                    265                    270  
 3  
 4    Glu Glu Ala Ala Lys Trp Gly Pro Leu Arg Trp Met Leu Ala Ile Gin  
 5                    275                    280                    285  
 6  
 7    Gln Met Thr Asn Ala Ala Ser Ser Tyr Met Thr Asp Tyr Tyr Ile Ser  
 8                    290                    295                    300  
 9  
 10   Thr Val Phe Gln Ala Arg His Ser Gln Asn Asn Tyr Leu Arg Val Gln  
 11                   305                    310                    315                    320  
 12  
 13   Glu Asn Ala Leu Asn Gly Thr Thr Thr Glu Met Asp Asp Ala Ser Glu  
 14                   325                    330                    335  
 15  
 16   Ala Asn Met Glu Leu Leu Val Gln Val Gly Ala Thr Leu Leu Lys Lys  
 17                   340                    345                    350  
 18  
 19   Pro Val Ser Lys Asp Ser Pro Glu Thr Tyr Glu Glu Ala Leu Lys Arg  
 20                   355                    360                    365  
 21  
 22   Phe Ala Lys Leu Leu Ser Asp Arg Lys Lys Leu Arg Ala Asn Lys Ala  
 23                   370                    375                    380  
 24  
 25   Ser Tyr  
 26                   385  
 27  
 28   <210> 283  
 29   <211> 10  
 30   <212> PRT  
 31   <213> Synthetic  
 32  
 33   <400> 283  
 34  
 35   Ala Phe Phe Asp Lys Thr Tyr Thr Ala Lys  
 36                   5    10  
 37  
 38   <210> 284  
 39   <211> 10  
 40   <212> PRT  
 41   <213> Synthetic  
 42  
 43   <400> 284  
 44  
 45   Cys Ile Phe Asp Ser Thr Tyr Thr Ala Lys  
 46                   5    10  
 47  
 48   <210> 285  
 49   <211> 1161  
 50   <212> DNA  
 51   <213> Solanum tuberosum  
 52  
 53   <400> 285  
 54   atggcaacta ctaaatcttt tttaatttta atatttatga tattagcaac tactagttca        60  
 55   acatttgctc agttgggaga aatggtgact gttcttagta ttgatggagg tggaaattaga        120  
 56

```

1 gggatcattc cggttacccat tctcgaattt cttgaaggac aacttcggaa aatggacaat 180
2 aatgcgatgg caagacttgc agattactt gatgttaattt gaggaaacaag tacaggaggt 240
3 ttattgtactg ctatgataag tactccaaat gaaaacaatc gaccctttgc tgctgc当地 300
4 gaaattgtac ctttttactt cgaacatggc cctcagattt ttaatcctag tggcaattt 360
5 ttagggccaa aatatgtgg aaaatattt atgcgaattt ttcaagaaaa acttggagaa 420
6 actcgtgtgc atcaagctt gacagaagtt gtcatctca gctttgacat caaaacaat 480
7 aaggccatggaa tatttactaa gtccaaattt gcaaaactctc cagaatttggaa tgctcaatg 540
8 tatgacataa gttattccac agcagcagttt ccaacatattt ttccctccgca ttatcttgg 600
9 actaataacta gtaatggaga tgaatatggat ttcaatcttgc ttgtatggc tggttgcact 660
10 gtgtgtgtatc cggcggttattt atccatttgc gtttgc当地 gacttgcaca aaaggatccca 720
11 gcatttgc当地 ctatgttgc attgaatttcc aaaaaatgc tggttgc当地 attaggactt 780
12 ggcactactt cagagtttgc taaaacatatac acagcaaaagg aggccagtttgc ctggactgt 840
13 gtacatggaa tgtagtttgc acagaaaaatgc actgtatggc caagtttgc当地 catgactgtat 900
14 tattaccctt ctactgttgc tcaaggttgc gattcaaaaaa acaatttgc当地 cagggttcaa 960
15 gaaaatgtatc taacaggccac aactacttgc当地 atggatgttgc ttccgtggcc taatatggaa 1020
16 ttattatgtatc aagtggatggaa aacttgc当地 aagaaaccatggaa ttccgtggcc taatccgttgc当地 1080
17 acctatgtatc aagtggatggaa aacttgc当地 aagaaaccatggaa ttccgtggcc taatccgttgc当地 1140
18 gcaaaacaatggaa ctttgc当地 a 1161
19 <210> 286
20 <211> 386
21 <212> PRT
22 <213> Solanum tuberosum
23
24 <400> 286
25
26 Met Ala Thr Thr Lys Ser Phe Leu Ile Leu Ile Phe Met Ile Leu Ala
27 1 5 10 15
28
29 Thr Thr Ser Ser Thr Phe Ala Gln Leu Gly Glu Met Val Thr Val Leu
30 20 25 30
31
32 Ser Ile Asp Gly Gly Gly Ile Arg Gly Ile Ile Pro Ala Thr Ile Leu
33 35 40 45
34 Glu Phe Leu Glu Gly Gln Leu Gln Glu Met Asp Asn Asn Ala Asp Ala
35 50 55 60
36 Arg Leu Ala Asp Tyr Phe Asp Val Ile Gly Gly Thr Ser Thr Gly Gly
37 65 70 75 80

```

1 Leu Leu Thr Ala Met Ile Ser Thr Pro Asn Glu Asn Asn Arg Pro Phe  
 2 85 90 95  
 3  
 4 Ala Ala Ala Lys Glu Ile Val Pro Phe Tyr Phe Glu His Gly Pro Gln  
 5 100 105 110  
 6  
 7 Ile Phe Asn Pro Ser Gly Gln Ile Leu Gly Pro Lys Tyr Asp Gly Lys  
 8 115 120 125  
 9  
 10 Tyr Leu Met Gln Val Leu Gln Glu Lys Leu Gly Glu Thr Arg Val His  
 11 130 135 140  
 12  
 13 Gln Ala Leu Thr Glu Val Val Ile Ser Ser Phe Asp Ile Lys Thr Asn  
 14 145 150 155 160  
 15  
 16 Lys Pro Val Ile Phe Thr Lys Ser Asn Leu Ala Asn Ser Pro Glu Leu  
 17 165 170 175  
 18  
 19 Asp Ala Lys Met Tyr Asp Ile Ser Tyr Ser Thr Ala Ala Ala Pro Thr  
 20 180 185 190  
 21  
 22 Tyr Phe Pro Pro His Tyr Phe Val Thr Asn Thr Ser Asn Gly Asp Glu  
 23 195 200 205  
 24  
 25 Tyr Glu Phe Asn Leu Val Asp Gly Ala Val Ala Thr Val Ala Asp Pro  
 26 210 215 220  
 27  
 28 Ala Leu Leu Ser Ile Ser Val Ala Thr Arg Leu Ala Gln Lys Asp Pro  
 29 225 230 235 240  
 30  
 31 Ala Phe Ala Ser Ile Arg Ser Leu Asn Tyr Lys Lys Met Leu Leu Leu  
 32 245 250 255  
 33  
 34 Ser Leu Gly Thr Gly Thr Thr Ser Glu Phe Asp Lys Thr Tyr Thr Ala  
 35 260 265 270  
 36  
 37 Lys Glu Ala Ala Thr Trp Thr Ala Val His Trp Met Leu Val Ile Gln  
 38 275 280 285  
 39  
 40 Lys Met Thr Asp Ala Ala Ser Ser Tyr Met Thr Asp Tyr Tyr Leu Ser  
 41 290 295 300  
 42  
 43 Thr Ala Phe Gln Ala Leu Asp Ser Lys Asn Asn Tyr Leu Arg Val Gln  
 44 305 310 315 320  
 45  
 46 Glu Asn Ala Leu Thr Gly Thr Thr Glu Met Asp Asp Ala Ser Glu  
 47 325 330 335  
 48  
 49 Ala Asn Met Glu Leu Leu Val Gln Val Gly Glu Asn Leu Leu Lys Lys  
 50 340 345 350  
 51  
 52 Pro Val Ser Glu Asp Asn Pro Glu Thr Tyr Glu Glu Ala Leu Lys Arg  
 53 355 360 365  
 54  
 55 Phe Ala Lys Leu Leu Ser Asp Arg Lys Lys Leu Arg Ala Asn Lys Ala  
 56 370 375 380  
 57

```

1 Ser Tyr
2 385
3
4
5 <210> 287
6 <211> 408
7 <212> PRT
8 <213> W098154327 12/3/98 (Equinous (Genus) Pentaclethaa species
9 macrolota)
10
11 <400> 287
12
13 Met Lys Ser Lys Met Ala Met Leu Leu Leu Leu Phe Cys Val Leu Ser
14 1 5 10 15
15
16 Asn Gln Leu Val Ala Ala Phe Ser Thr Gln Ala Lys Ala Ser Lys Asp
17 20 25 30
18
19 Gly Asn Leu Val Thr Val Leu Ala Ile Asp Gly Gly Ile Arg Gly
20 35 40 45
21
22 Ile Ile Pro Gly Val Ile Leu Lys Gln Leu Glu Ala Thr Leu Gln Arg
23 50 55 60
24
25 Trp Asp Ser Ser Ala Arg Leu Ala Glu Tyr Phe Asp Val Val Ala Gly
26 65 70 75 80
27
28 Thr Ser Thr Gly Gly Ile Ile Thr Ala Ile Leu Thr Ala Pro Asp Pro
29 85 90 95
30
31 Gln Asn Lys Asp Arg Pro Leu Tyr Ala Ala Glu Glu Ile Ile Asp Phe
32 100 105 110
33
34 Tyr Ile Glu His Gly Pro Ser Ile Phe Asn Lys Ser Thr Ala Cys Ser
35 115 120 125
36
37 Leu Pro Gly Ile Phe Cys Pro Lys Tyr Asp Gly Lys Tyr Leu Gln Glu
38 130 135 140
39
40 Ile Ile Ser Gln Lys Leu Asn Glu Thr Leu Leu Asp Gln Thr Thr Thr
41 145 150 155 160
42
43 Asn Val Val Ile Pro Ser Phe Asp Ile Lys Leu Leu Arg Pro Thr Ile
44 165 170 175
45
46 Phe Ser Thr Phe Lys Leu Glu Glu Val Pro Glu Leu Asn Val Lys Leu
47 180 185 190
48
49 Ser Asp Val Cys Met Gly Thr Ser Ala Ala Pro Ile Val Phe Pro Pro
50 195 200 205
51
52 Tyr Tyr Phe Lys His Gly Asp Thr Glu Phe Asn Leu Val Asp Gly Ala
53 210 215 220
54
55 Ile Ile Ala Asp Ile Pro Ala Pro Val Ala Leu Ser Glu Val Leu Gln
56 225 230 235
57

```

1 Gln Glu Lys Tyr Lys Asn Lys Glu Ile Leu Leu Leu Ser Ile Gly Thr  
2 245 250 255  
3  
4 Gly Val Val Lys Pro Gly Glu Gly Tyr Ser Ala Asn Arg Thr Trp Thr  
5 260 265 270  
6  
7 Ile Phe Asp Trp Ser Ser Glu Thr Leu Ile Gly Leu Met Gly His Gly  
8 275 280 285  
9  
10 Thr Arg Ala Met Ser Asp Tyr Tyr Val Gly Ser His Phe Lys Ala Leu  
11 290 295 300  
12  
13 Gln Pro Gln Asn Asn Tyr Leu Arg Ile Gln Glu Tyr Asp Leu Asp Pro  
14 305 310 315 320  
15  
16 Ala Leu Glu Ser Ile Asp Asp Ala Ser Thr Glu Asn Met Glu Asn Leu  
17 325 330 335  
18  
19 Glu Lys Val Gly Gln Ser Leu Leu Asn Glu Pro Val Lys Arg Met Asn  
20 340 345 350  
21  
22 Leu Asn Thr Phe Val Val Glu Glu Thr Gly Glu Gly Thr Asn Ala Glu  
23 355 360 365  
24  
25 Ala Leu Asp Arg Leu Ala Gln Ile Leu Tyr Glu Glu Lys Ile Thr Arg  
26 370 375 380  
27  
28 Gly Leu Gly Lys Ile Ser Leu Glu Val Asp Asn Ile Asp Pro Tyr Thr  
29 385 390 395 400  
30  
31 Glu Arg Val Arg Lys Leu Leu Phe  
32 405  
33  
34 <210> 288  
35 <211> 410  
36 <212> PRT  
37 <213> Zea mays  
38  
39 <400> 288  
40  
41 Met Gly Ser Ile Gly Arg Gly Thr Ala Asn Cys Ala Thr Val Pro Gln  
42 1 5 10 15  
43  
44 Pro Pro Pro Ser Thr Gly Lys Leu Ile Thr Ile Leu Ser Ile Asp Gly  
45 20 25 30  
46  
47 Gly Gly Ile Arg Gly Leu Ile Pro Ala Thr Ile Ile Ala Tyr Leu Glu  
48 35 40 45  
49  
50 Ala Lys Leu Gln Glu Leu Asp Gly Pro Asp Ala Arg Ile Ala Asp Tyr  
51 50 55 60  
52  
53 Phe Asp Val Ile Ala Gly Thr Ser Thr Gly Ala Leu Leu Ala Ser Met  
54 65 70 75 80  
55  
56 Leu Ala Ala Pro Asp Glu Asn Asn Arg Pro Leu Phe Ala Ala Lys Asp  
57 85 90 95

1 Leu Thr Thr Phe Tyr Leu Glu Asn Gly Pro Lys Ile Phe Pro Gln Lys  
2 100 105 110  
3  
4 Lys Ala Gly Leu Leu Thr Pro Leu Arg Asn Leu Leu Gly Leu Val Arg  
5 115 120 125  
6  
7 Gly Pro Lys Tyr Asp Gly Val Phe Leu His Asp Lys Ile Lys Ser Leu  
8 130 135 140  
9  
10 Thr His Asp Val Arg Val Ala Asp Thr Val Thr Asn Val Ile Val Pro  
11 145 150 155 160  
12  
13 Ala Phe Asp Val Lys Tyr Leu Gln Pro Ile Ile Phe Ser Thr Tyr Glu  
14 165 170 175  
15  
16 Ala Lys Thr Asp Thr Leu Lys Asn Ala His Leu Ser Asp Ile Cys Ile  
17 180 185 190  
18  
19 Ser Thr Ser Ala Ala Pro Thr Tyr Phe Pro Ala His Phe Phe Lys Thr  
20 195 200 205  
21  
22 Glu Ala Thr Asp Gly Arg Pro Pro Arg Glu Tyr His Leu Val Asp Gly  
23 210 215 220  
24  
25 Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Ser Met Leu Thr  
26 225 230 235 240  
27  
28 Lys Glu Val His Arg Arg Asn Pro Asn Phe Asn Ala Gly Ser Pro Thr  
29 245 250 255  
30  
31 Glu Tyr Thr Asn Tyr Leu Ile Ile Ser Val Gly Thr Gly Ser Ala Lys  
32 260 265 270  
33  
34 Gln Ala Glu Lys Tyr Thr Ala Glu Gln Cys Ala Lys Trp Gly Leu Ile  
35 275 280 285  
36  
37 Gln Trp Leu Tyr Asn Gly Gly Phe Thr Pro Ile Ile Asp Ile Phe Ser  
38 290 295 300  
39  
40 His Ala Ser Ser Asp Met Val Asp Ile His Ala Ser Ile Leu Phe Gln  
41 305 310 315 320  
42  
43 Ala Leu His Cys Glu Lys Lys Tyr Leu Arg Ile Gln Asp Asp Thr Leu  
44 325 330 335  
45  
46 Thr Gly Asn Ala Ser Ser Val Asp Ile Ala Thr Lys Glu Asn Met Glu  
47 340 345 350  
48  
49 Ser Leu Ile Ser Ile Gly Gln Glu Leu Leu Lys Lys Pro Val Ala Arg  
50 355 360 365  
51  
52 Val Asn Ile Asp Thr Gly Val Tyr Glu Ser Cys Asp Gly Glu Gly Thr  
53 370 375 380  
54  
55 Asn Ala Gln Ser Leu Ala Asp Phe Ala Lys Gln Leu Ser Asp Glu Arg  
56 385 390 395 400  
57

1 Lys Leu Arg Lys Ser Asn Leu Asn Ser Asn  
2 405 410  
3  
4  
5 <210> 289  
6 <211> 508  
7 <212> PRT  
8 <213> Zea mays  
9  
10 <400> 289  
11  
12 Arg Pro Thr Arg Pro Arg His Pro Arg Asn Thr Gln Lys Arg Gly Ala  
13 5 10 15  
14  
15 Leu Leu Val Gly Trp Ile Leu Phe Ser Leu Ala Ala Ser Pro Val Lys  
16 20 25 30  
17  
18 Phe Gln Thr His Met Gly Ser Ile Gly Arg Gly Thr Ala Asn Cys Ala  
19 35 40 45  
20  
21 Thr Val Pro Gln Pro Pro Ser Thr Gly Lys Leu Ile Thr Ile Leu  
22 50 55 60  
23  
24 Ser Ile Asp Gly Gly Ile Arg Gly Leu Ile Pro Ala Thr Ile Ile  
25 65 70 75 80  
26  
27 Ala Tyr Leu Glu Ala Lys Leu Gln Glu Leu Asp Gly Pro Asp Ala Arg  
28 85 90 95  
29  
30 Ile Ala Asp Tyr Phe Asp Val Ile Ala Gly Thr Ser Thr Gly Ala Leu  
31 100 105 110  
32  
33 Leu Ala Ser Met Leu Ala Ala Pro Asp Glu Asn Asn Arg Pro Leu Phe  
34 115 120 125  
35  
36 Ala Ala Lys Asp Leu Thr Thr Phe Tyr Leu Glu Asn Gly Pro Lys Ile  
37 130 135 140  
38  
39 Phe Pro Gln Lys Lys Ala Gly Leu Leu Thr Pro Leu Arg Asn Leu Leu  
40 145 150 155 160  
41  
42 Gly Leu Val Arg Gly Pro Lys Tyr Asp Gly Val Phe Leu His Asp Lys  
43 165 170 175  
44  
45 Ile Lys Ser Leu Thr His Asp Val Arg Val Ala Asp Thr Val Thr Asn  
46 180 185 190  
47  
48 Val Ile Val Pro Ala Phe Asp Val Lys Tyr Leu Gln Pro Ile Ile Phe  
49 195 200 205  
50  
51 Ser Thr Tyr Glu Ala Lys Thr Asp Ala Leu Lys Asn Ala His Leu Ser  
52 210 215 220  
53  
54 Asp Ile Cys Ile Ser Thr Ser Ala Ala Pro Thr Tyr Phe Pro Ala His  
55 225 230 235 240  
56  
57 Phe Phe Lys Thr Glu Ala Thr Asp Gly Arg Pro Pro Arg Glu Tyr His

1	245	250	255
2	Leu Val Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met		
3	260	265	270
4			
5	Ser Met Leu Thr Lys Glu Val His Arg Arg Asn Pro Asn Phe Asn Ala		
6	275	280	285
7			
8	Gly Ser Pro Thr Glu Tyr Thr Asn Tyr Leu Ile Ile Ser Val Gly Thr		
9	290	295	300
10			
11	Gly Ser Ala Lys Gln Ala Glu Lys Tyr Thr Ala Glu Gln Cys Ala Lys		
12	305	310	315
13			320
14			
15	Trp Gly Leu Ile Gln Trp Leu Tyr Asn Gly Gly Phe Thr Pro Ile Ile		
16	325	330	335
17			
18	Asp Ile Phe Ser His Ala Ser Ser Asp Met Val Asp Ile His Ala Ser		
19	340	345	350
20			
21	Ile Leu Phe Gln Ala Leu His Cys Glu Lys Lys Tyr Leu Arg Ile Gln		
22	355	360	365
23			
24	Leu Tyr Tyr Ala Gly Tyr Phe Asp Trp Glu Arg Ile Val Arg Gly His		
25	370	375	380
26			
27	Arg His Gln Gly Glu His Gly Val Ser Asp Ile Asp Arg Pro Gly Ala		
28	385	390	395
29			400
30	Ala Gln Glu Ala Ser Gly Glu Ser Glu His Arg His Arg Ala Val Arg		
31	405	410	415
32			
33	Val Leu Arg Arg Gly His Lys Cys Thr Val Ala Ser Leu Arg Gln Ala		
34	420	425	430
35			
36	Thr Leu Arg Ala Gln Ala Thr Gln Glu Gln Ser Gln Leu Gln Leu Ile		
37	435	440	445
38			
39	Asn Thr Ser Leu Ser His Ser Met Cys Ser Phe Arg Arg Phe Thr Val		
40	450	455	460
41			
42	Ser Tyr Phe Phe Asn Phe Asn Ser Val Cys Val Leu Cys Val Leu Cys		
43	465	470	475
44			480
45	Val Tyr Gln Thr Phe Lys Phe Asn Gln Lys Lys Lys Lys Lys Lys Lys		
46	485	490	495
47			
48	Lys Lys Lys Lys Lys Lys Lys Lys Arg Ala Ala		
49	500	505	
50			
51	<210> 290		
52	<211> 410		
53	<212> PRT		
54	<213> Zea mays		
55			
56	<400> 290		
57			

1 Met Gly Ser Ile Gly Arg Gly Thr Ala Asn Cys Ala Thr Val Pro Gln  
 2 1 5 10 15  
 3  
 4 Pro Pro Pro Ser Thr Gly Lys Leu Ile Thr Ile Leu Ser Ile Asp Gly  
 5 20 25 30  
 6  
 7 Gly Gly Ile Arg Gly Leu Ile Pro Ala Thr Ile Ile Ala Tyr Leu Glu  
 8 35 40 45  
 9  
 10 Ala Lys Leu Gln Glu Leu Asp Gly Pro Asp Ala Arg Ile Ala Asp Tyr  
 11 50 55 60  
 12  
 13 Phe Asp Val Ile Ala Gly Thr Ser Thr Gly Ala Leu Leu Ala Ser Met  
 14 65 70 75 80  
 15  
 16 Leu Ala Ala Pro Asp Glu Asn Asn Arg Pro Leu Phe Ala Ala Lys Asp  
 17 85 90 95  
 18  
 19 Leu Thr Thr Phe Tyr Leu Glu Asn Gly Pro Lys Ile Phe Pro Gln Lys  
 20 100 105 110  
 21  
 22 Lys Ala Gly Leu Leu Thr Pro Leu Arg Asn Leu Leu Gly Leu Val Arg  
 23 115 120 125  
 24  
 25 Gly Pro Lys Tyr Asp Gly Val Phe Leu His Asp Lys Ile Lys Ser Leu  
 26 130 135 140  
 27  
 28 Thr His Asp Val Arg Val Ala Asp Thr Val Thr Asn Val Ile Val Pro  
 29 145 150 155 160  
 30  
 31 Ala Phe Asp Val Lys Ser Leu Gln Pro Ile Ile Phe Ser Thr Tyr Glu  
 32 165 170 175  
 33  
 34 Ala Lys Thr Asp Thr Leu Lys Asn Ala His Leu Ser Asp Ile Cys Ile  
 35 180 185 190  
 36  
 37 Ser Thr Ser Ala Ala Pro Thr Tyr Phe Pro Ala His Phe Phe Lys Thr  
 38 195 200 205  
 39  
 40 Glu Ala Thr Asp Gly Arg Pro Pro Arg Glu Tyr His Leu Val Asp Gly  
 41 210 215 220  
 42  
 43 Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Ser Met Leu Thr  
 44 225 230 235 240  
 45  
 46 Lys Glu Val His Arg Arg Asn Pro Asn Phe Asn Ala Gly Ser Pro Thr  
 47 245 250 255  
 48  
 49 Glu Tyr Thr Asn Tyr Leu Ile Ile Ser Val Gly Thr Gly Ser Ala Lys  
 50 260 265 270  
 51  
 52 Gln Ala Glu Lys Tyr Thr Ala Glu Gln Cys Ala Lys Trp Gly Leu Ile  
 53 275 280 285  
 54  
 55 Gln Trp Leu Tyr Asn Gly Gly Phe Thr Pro Ile Ile Asp Ile Phe Ser  
 56 290 295 300  
 57

1 His Ala Ser Ser Asp Met Val Asp Ile His Ala Ser Ile Leu Phe Gln  
 2 305 310 315 320  
 3  
 4 Ala Leu His Cys Glu Lys Lys Tyr Leu Arg Ile Gln Asp Asp Thr Leu  
 5 325 330 335  
 6  
 7 Thr Gly Asn Ala Ser Ser Val Asp Ile Ala Thr Lys Glu Asn Met Glu  
 8 340 345 350  
 9  
 10 Ser Leu Ile Ser Ile Gly Gln Glu Leu Leu Asn Lys Pro Val Ala Arg  
 11 355 360 365  
 12  
 13 Val Asn Ile Asp Thr Gly Leu Tyr Glu Ser Cys Glu Gly Glu Gly Thr  
 14 370 375 380  
 15  
 16 Asn Ala Gln Ser Leu Ala Asp Phe Ala Lys Gln Leu Ser Asp Glu Arg  
 17 385 390 395 400  
 18  
 19 Lys Leu Arg Lys Ser Asn Leu Asn Ser Asn  
 20 405 410  
 21  
 22 <210> 291  
 23 <211> 410  
 24 <212> PRT  
 25 <213> Zea mays  
 26  
 27 <400> 291  
 28  
 29 Met Gly Ser Ile Gly Arg Gly Thr Ala Asn Cys Ala Thr Val Pro Gln  
 30 1 5 10 15  
 31  
 32 Pro Pro Pro Ser Thr Gly Lys Leu Ile Thr Ile Leu Ser Ile Asp Gly  
 33 20 25 30  
 34  
 35 Gly Gly Ile Arg Gly Leu Ile Pro Ala Thr Ile Ile Ala Tyr Leu Glu  
 36 35 40 45  
 37  
 38 Ala Lys Leu Gln Glu Leu Asp Gly Pro Asp Ala Arg Ile Ala Asp Tyr  
 39 50 55 60  
 40  
 41 Phe Asp Val Ile Ala Gly Thr Ser Thr Gly Ala Leu Leu Ala Ser Met  
 42 65 70 75 80  
 43  
 44 Leu Ala Ala Pro Asp Glu Asn Asn Arg Pro Leu Phe Ala Ala Lys Asp  
 45 85 90 95  
 46  
 47 Leu Thr Thr Phe Tyr Leu Glu Asn Gly Pro Lys Ile Phe Pro Gln Lys  
 48 100 105 110  
 49  
 50 Lys Ala Gly Leu Leu Thr Pro Leu Arg Asn Leu Leu Gly Leu Val Arg  
 51 115 120 125  
 52  
 53 Gly Pro Lys Tyr Asp Gly Val Phe Leu His Asp Lys Ile Lys Ser Leu  
 54 130 135 140  
 55  
 56 Thr His Asp Val Arg Val Ala Asp Thr Val Thr Asn Val Ile Val Pro  
 57 145 150 155 160

1 Ala Phe Asp Val Lys Tyr Leu Gln Pro Ile Ile Phe Ser Thr Tyr Glu  
2 165 170 175  
3  
4 Ala Lys Thr Asp Ala Leu Lys Asn Ala His Leu Ser Asp Ile Cys Ile  
5 180 185 190  
6  
7 Ser Thr Ser Ala Ala Pro Thr Tyr Phe Pro Ala His Phe Phe Lys Thr  
8 195 200 205  
9  
10 Glu Ala Thr Asp Gly Arg Pro Pro Arg Glu Tyr His Leu Val Asp Gly  
11 210 215 220  
12  
13 Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Ser Met Leu Thr  
14 225 230 235 240  
15  
16 Lys Glu Val His Arg Arg Asn Pro Asn Phe Asn Ala Gly Ser Pro Thr  
17 245 250 255  
18  
19 Glu Tyr Thr Asn Tyr Leu Ile Ile Ser Val Gly Thr Gly Ser Ala Lys  
20 260 265 270  
21  
22 Gln Ala Glu Lys Tyr Thr Ala Glu Gln Cys Ala Lys Trp Gly Leu Ile  
23 275 280 285  
24  
25 Gln Trp Leu Tyr Asn Gly Gly Phe Thr Pro Ile Ile Asp Ile Phe Ser  
26 290 295 300  
27  
28 His Ala Ser Ser Asp Met Val Asp Ile His Ala Ser Ile Leu Phe Gln  
29 305 310 315 320  
30  
31 Ala Leu His Cys Glu Lys Lys Tyr Leu Arg Ile Gln Asp Asp Thr Leu  
32 325 330 335  
33  
34 Thr Gly Asn Ala Ser Ser Val Asp Ile Ala Thr Lys Glu Asn Met Glu  
35 340 345 350  
36  
37 Ser Leu Ile Ser Ile Gly Gln Glu Leu Leu Lys Lys Pro Val Ala Arg  
38 355 360 365  
39  
40 Val Asn Ile Asp Thr Gly Leu Tyr Glu Ser Cys Asp Gly Glu Gly Thr  
41 370 375 380  
42  
43 Asn Ala Gln Ser Leu Ala Asp Phe Ala Lys Gln Leu Ser Asp Glu Arg  
44 385 390 395 400  
45  
46 Lys Leu Arg Lys Ser Asn Leu Asn Ser Asn  
47 405 410  
48  
49 .  
50 <210> 292  
51 <211> 410  
52 <212> PRT  
53 <213> Zea mays  
54  
55 <400> 292  
56  
57 Met Gly Ser Ile Gly Arg Gly Thr Ala Asn Cys Ala Thr Val Pro Gln

1	1	5	10	15
2				
3	Pro Pro Pro Ser Thr Gly Lys Leu Ile Thr Ile Leu Ser Ile Asp Gly			
4	20	25	30	
5				
6	Gly Gly Ile Arg Gly Leu Ile Pro Ala Thr Ile Ile Ala Tyr Leu Glu			
7	35	40	45	
8				
9	Ala Lys Leu Gln Glu Leu Asp Gly Pro Asp Ala Arg Ile Ala Asp Tyr			
10	50	55	60	
11				
12	Phe Asp Val Ile Ala Gly Thr Ser Thr Gly Ala Leu Leu Ala Ser Met			
13	65	70	75	80
14				
15	Leu Ala Ala Pro Asp Glu Asn Asn Arg Pro Leu Phe Ala Ala Lys Asp			
16	85	90	95	
17				
18	Leu Thr Thr Phe Tyr Leu Glu Asn Gly Pro Lys Ile Phe Pro Gln Lys			
19	100	105	110	
20				
21	Lys Ala Gly Leu Leu Thr Pro Leu Arg Asn Leu Leu Gly Leu Val Arg			
22	115	120	125	
23				
24	Gly Pro Lys Tyr Asp Gly Val Phe Leu His Asp Lys Ile Lys Ser Leu			
25	130	135	140	
26				
27	Thr His Asp Val Arg Val Ala Asp Thr Val Thr Asn Val Ile Val Pro			
28	145	150	155	160
29				
30	Ala Phe Asp Val Lys Ser Leu Gln Pro Ile Ile Phe Ser Thr Tyr Glu			
31	165	170	175	
32				
33	Ala Lys Thr Asp Thr Leu Lys Asn Ala His Leu Ser Asp Ile Cys Ile			
34	180	185	190	
35				
36	Ser Thr Ser Ala Ala Pro Thr Tyr Phe Pro Ala His Phe Phe Lys Ile			
37	195	200	205	
38				
39	Glu Ala Thr Asp Gly Arg Pro Pro Arg Glu Tyr His Leu Val Asp Gly			
40	210	215	220	
41				
42	Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Ser Met Leu Thr			
43	225	230	235	240
44				
45	Lys Glu Val His Arg Arg Asn Pro Asn Phe Asn Ala Gly Ser Pro Thr			
46	245	250	255	
47				
48	Glu Tyr Thr Asn Tyr Leu Ile Ile Ser Val Gly Thr Gly Ser Ala Lys			
49	260	265	270	
50				
51	Gln Ala Glu Lys Tyr Thr Ala Glu Gln Cys Ala Lys Trp Gly Leu Ile			
52	275	280	285	
53				
54	Gln Trp Leu Tyr Asn Gly Gly Phe Thr Pro Ile Ile Asp Ile Phe Ser			
55	290	295	300	
56				
57	His Ala Ser Ser Asp Met Val Asp Ile His Ala Ser Ile Leu Phe Gln			

1	305	310	315	320
2				
3	Ala Leu His Cys Glu Lys Lys Tyr Leu Arg Ile Gln Asp Asp Thr Leu			
4	325	330	335	
5				
6	Thr Gly Asn Ala Ser Ser Val Asp Ile Ala Thr Lys Glu Asn Met Glu			
7	340	345	350	
8				
9	Ser Leu Ile Ser Ile Gly Gln Glu Leu Leu Asn Lys Pro Val Ala Arg			
10	355	360	365	
11				
12	Val Asn Ile Asp Thr Gly Leu Tyr Glu Ser Cys Glu Gly Glu Gly Thr			
13	370	375	380	
14				
15	Asn Ala Gln Ser Leu Ala Asp Phe Ala Lys Gln Leu Ser Asp Glu Arg			
16	385	390	395	400
17				
18	Lys Leu Arg Lys Ser Asn Leu Asn Ser Asn			
19	405	410		
20				
21	<210> 293			
22	<211> 337			
23	<212> PRT			
24	<213> Zea mays			
25				
26	<400> 293			
27				
28	Met Gly Ser Ile Gly Arg Gly Thr Ala Asn Cys Ala Thr Val Pro Gln			
29	1 5 10 15			
30				
31	Pro Pro Pro Ser Thr Gly Lys Leu Ile Thr Ile Leu Ser Ile Asp Gly			
32	20 25 30			
33				
34	Gly Gly Ile Arg Gly Leu Ile Pro Ala Thr Ile Ile Ala Tyr Leu Glu			
35	35 40 45			
36				
37	Ala Lys Leu Gln Glu Leu Asp Gly Pro Asp Ala Arg Ile Ala Asp Tyr			
38	50 55 60			
39				
40	Phe Asp Val Ile Ala Gly Thr Ser Thr Gly Ala Leu Leu Ala Ser Met			
41	65 70 75 80			
42				
43	Leu Ala Ala Pro Asp Glu Asn Asn Arg Pro Leu Phe Ala Ala Lys Asp			
44	85 90 95			
45				
46	Leu Thr Thr Phe Tyr Leu Glu Asn Gly Pro Lys Ile Phe Pro Gln Lys			
47	100 105 110			
48				
49	Lys Ala Gly Leu Leu Thr Pro Leu Arg Asn Leu Leu Gly Leu Val Arg			
50	115 120 125			
51				
52	Gly Pro Lys Tyr Asp Gly Val Phe Leu His Asp Lys Ile Lys Ser Leu			
53	130 135 140			
54				
55	Thr His Asp Val Arg Val Ala Asp Thr Val Thr Asn Val Ile Val Pro			
56	145 150 155			160
57				

1 Ala Phe Asp Val Lys Tyr Leu Gln Pro Ile Ile Phe Ser Thr Tyr Glu  
2 165 170 175  
3  
4 Ala Lys Thr Asp Ala Leu Lys Asn Ala His Leu Ser Asp Ile Cys Ile  
5 180 185 190  
6  
7 Ser Thr Ser Ala Ala Pro Thr Tyr Phe Pro Ala His Phe Phe Lys Thr  
8 195 200 205  
9  
10 Glu Ala Thr Asp Gly Arg Pro Pro Arg Glu Tyr His Leu Val Asp Gly  
11 210 215 220  
12  
13 Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Ser Met Leu Thr  
14 225 230 235 240  
15  
16 Lys Glu Val His Arg Arg Asn Pro Asn Phe Asn Ala Gly Ser Pro Thr  
17 245 250 255  
18  
19 Glu Tyr Thr Asn Tyr Leu Ile Ile Ser Val Gly Thr Gly Ser Ala Lys  
20 260 265 270  
21  
22 Gln Ala Glu Lys Tyr Thr Ala Glu Gln Cys Ala Lys Trp Gly Leu Ile  
23 275 280 285  
24  
25 Gln Trp Leu Tyr Asn Gly Gly Phe Thr Pro Ile Ile Asp Ile Phe Ser  
26 290 295 300  
27  
28 His Ala Ser Ser Asp Met Val Asp Ile His Ala Ser Ile Leu Phe Gln  
29 305 310 315 320  
30  
31 Ala Leu His Cys Glu Lys Lys Tyr Leu Arg Ile Gln Leu Tyr Tyr Ala  
32 325 330 335  
33  
34 Gly  
35  
36  
37  
38  
39